



OPTICAL FLOW-BASED ODOMETRY  
FOR  
UNDERGROUND TUNNEL EXPLORATION

THESIS

Terra Kier, First Lieutenant, USAF

AFIT/GE/ENG/11-21

DEPARTMENT OF THE AIR FORCE  
AIR UNIVERSITY

**AIR FORCE INSTITUTE OF TECHNOLOGY**

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States

AFIT/GE/ENG/11-21

OPTICAL FLOW-BASED ODOMETRY  
FOR  
UNDERGROUND TUNNEL EXPLORATION

THESIS

Presented to the Faculty  
Department of Electrical and Computer Engineering  
Graduate School of Engineering and Management  
Air Force Institute of Technology  
Air University  
Air Education and Training Command  
In Partial Fulfillment of the Requirements for the  
Degree of Master of Science in Electrical Engineering

Terra Kier, B.S.E.E.  
First Lieutenant, USAF

March 2011


APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT/GE/ENG/11-21

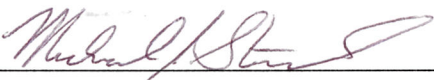
OPTICAL FLOW-BASED ODOMETRY  
FOR  
UNDERGROUND TUNNEL EXPLORATION

Terra Kier, B.S.E.E.  
First Lieutenant, USAF

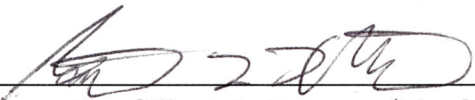
Approved:

  
Maj Kenneth A. Fisher, PhD (Chairman)

23 March 2011  
date

  
Lt Col Michael J. Stepaniak (Member)

23 Mar 2011  
date

  
Dr. Gilbert L. Peterson (Member)

23 MAR 2011  
date

*Abstract*

As military operations in degraded or GPS-denied environments continue to increase in frequency and importance, there is an increased necessity to be able to determine precision location within these environments. Furthermore, authorities are finding a record number of tunnels along the U.S.-Mexico border; therefore, underground tunnel characterization is becoming a high priority for U.S. Homeland Security as well. This thesis investigates the performance of a new image registration technique based on a two camera optical-flow configuration using phase correlation techniques. These techniques differ from other image based navigation methods but present a viable alternative increasing autonomy and answering the tunnel based navigation problem. This research presents an optical flow based image registration technique and validates its use with experimental results on a mobile vehicle. Results reveal that further extension to pose estimation yields accuracy within 1.3 cm.

## *Acknowledgements*

I would like to express my sincere appreciation to my faculty advisor, Major Ken Fisher, for his guidance and support throughout the course of this thesis effort. The insight, direction, and advice were certainly appreciated. I would also like to thank Mr. Don Smith, as well as Mr. Fred and Mr. Glasses, for their assistance and experiment setup support. I would also like to thank my parents for all the sacrifices that they have made to help me get where I am today and for instilling in me the character, dedication, and work ethic required to complete tasks such as these.

Terra Kier

# Table of Contents

	Page
Abstract . . . . .	iv
Acknowledgements . . . . .	v
List of Figures . . . . .	viii
List of Tables . . . . .	x
List of Abbreviations . . . . .	xi
I. Introduction . . . . .	1
1.1 Current Technology . . . . .	1
1.2 Scope . . . . .	3
1.3 Organization . . . . .	3
II. Background . . . . .	4
2.1 Mathematical Background . . . . .	4
2.1.1 Coordinate Systems and Transformations . . . . .	4
2.1.2 Coordinate Transformations . . . . .	6
2.2 Image Registration . . . . .	9
2.2.1 Digital Image Processing . . . . .	9
2.2.2 Cross-Correlation . . . . .	9
2.2.3 Phase Correlation . . . . .	10
2.3 Non-GPS Navigation Techniques . . . . .	14
2.3.1 Dead Reckoning . . . . .	16
2.3.2 Navigation Based on <i>A Priori</i> Maps . . . . .	16
2.3.3 Laser-Based Navigation . . . . .	16
2.3.4 Signals of Opportunity Navigation . . . . .	17
2.3.5 Visual Based Navigation . . . . .	17
2.4 Summary . . . . .	18
III. Methodology . . . . .	19
3.1 Theory and Algorithms . . . . .	19
3.1.1 Phase One: Single Image Characterization . . . . .	19
3.1.2 Phase Two: Single Camera Characterization . . . . .	21
3.1.3 Phase Correlation . . . . .	22
3.1.4 Phase Three: Two Camera Characterization . . . . .	23
3.2 Performance Metrics . . . . .	29
3.3 Summary . . . . .	31

	Page
IV. Analysis and Results . . . . .	34
4.1 Hardware Overview . . . . .	34
4.1.1 Digital Cameras . . . . .	34
4.1.2 Tactical Grade IMU . . . . .	34
4.1.3 SPAN . . . . .	35
4.1.4 Vehicle Platform . . . . .	35
4.2 Simulation Results . . . . .	36
4.2.1 Peak Parameter Sweep . . . . .	37
4.2.2 Displacement Error Calculation . . . . .	39
4.3 Summary . . . . .	42
V. Conclusions . . . . .	45
5.1 Conclusions . . . . .	45
5.2 Recommendations for Future Research . . . . .	46
Bibliography . . . . .	48



## *List of Figures*

Figure		Page
1.1.	Surge in Tunnel Excavations . . . . .	2
2.1.	Earth-Centered Earth-Fixed Frame. . . . .	6
2.2.	Earth-fixed Local-Level Navigation Frame. . . . .	7
2.3.	Image Shift: $I_2$ is a Shifted Copy of $I_1$ . . . . .	11
2.4.	Periodic Extensions Due to Fourier Methods of Cross-Correlation. . . . .	12
2.5.	Zero-Padded Image To Prevent the Wraparound Effect. . . . .	13
2.6.	Cross-Correlation Plot Based on Fourier Methods. . . . .	14
2.7.	Curvature of the Cross-Correlation Plot. . . . .	15
3.1.	Optical Flow Based Registration Algorithm. . . . .	20
3.2.	Single Large Image. . . . .	21
3.3.	Image Crop 1. . . . .	22
3.4.	Image Crop 2. . . . .	23
3.5.	Single Image Cross-Correlation. . . . .	24
3.6.	Determination of Shift Example. . . . .	25
3.7.	Single Camera Cross-Correlation. . . . .	26
3.8.	Single Camera Curvature of the Cross-Correlation. . . . .	27
3.9.	Single Camera Cross-Correlation No Correlation Determined. . . . .	28
3.10.	Single Camera Curvature of the Cross-Correlation No Correlation Determined. . . . .	29
3.11.	Optical Flow From Single Camera Image Processing. . . . .	30
3.12.	Image Captures From Single Camera Image Processing. . . . .	31
3.13.	Cross-Correlation from Single Camera Image Processing. . . . .	31
3.14.	Positioning of the Cameras in the Body Frame. . . . .	32
3.15.	Two different Paths That Return the Same Readings. . . . .	32
3.16.	Triangle Used to Resolve $\Delta\theta$ Based on the Cosine Rule. . . . .	33

Figure		Page
4.1.	Implementation Block Diagram. . . . .	35
4.2.	Vehicle platform setup. . . . .	36
4.3.	Vehicle platform setup. . . . .	37
4.4.	Alpha Parameter Sweep For X-Direction Test Setup. . . . .	38
4.5.	Alpha Parameter Sweep For Y-Direction Test Setup. . . . .	39
4.6.	Alpha Parameter Sweep For Diagonal Test Setup. . . . .	40
4.7.	Alpha Parameter Sweep Data. . . . .	41
4.8.	Statistics for X-Displacement Differences. . . . .	42
4.9.	Statistics for Y-Displacement Differences. . . . .	43
4.10.	Average Velocity Over Test Run . . . . .	43
4.11.	Multidimensional Displacement Error . . . . .	44

# *List of Tables*

Table		Page
4.1.	Statistics for X-Displacement Differences. . . . .	39
4.2.	Statistics for X and Y-Displacement Differences. . . . .	40
4.3.	Statistics for Multidimensional Displacement Differences. . . .	41

# *List of Abbreviations*

Abbreviation		Page
GPS	Global Positioning System . . . . .	1
NED	North East Down . . . . .	6
DCM	Direction Cosine Matrices . . . . .	6
FFT	Fast Fourier Transform . . . . .	9
PCM	Phase Correlation Method . . . . .	10
SoOP	Signals of Opportunity . . . . .	15
TDOA	Time Difference of Arrival . . . . .	17
AM	Amplitude Modulation . . . . .	17
SLAM	Simultaneous Localization and Mapping . . . . .	17
IMU	Inertial Measurement Unit . . . . .	34
SPAN	Synchronous Position, Attitude and Navigation . . . . .	34

# OPTICAL FLOW-BASED ODOMETRY FOR UNDERGROUND TUNNEL EXPLORATION

## I. Introduction

This thesis describes the development of a robust optical flow based image registration technique for navigation aiding in Global Positioning System (GPS) denied environments. This research is motivated by the requirement for precision navigation in uncharacterized tunnel environments.

Authorities are finding a record number of tunnels along the U.S.-Mexico border, which is alarming for U.S. security as smugglers are becoming more ingenuitive in their drug and human smuggling tactics. Since the Secure Border Fence Act passed in 2006, tunnel discoveries have skyrocketed as shown in Figure 1.1 [18].

Determining the origin and destination of unknown tunnels is gaining increasing importance. GPS is viable only when there is an unobstructed view of the sky to receive the satellite signals, making it unavailable for underground navigation. Robotic platforms used for reconnaissance in the unknown tunnel environment will not only be a valuable asset to GPS denied areas, but a necessity for future operations.

### ***1.1 Current Technology***

Researchers have implemented many navigation techniques to solve the robotic navigation problem [3]. However, most of these methods do not allow for extension into the underground navigation problem [13]. GPS-navigation has been a proven method for autonomous navigation only when an unobstructed view of the sky is available, therefore this is not a viable option. Signals of opportunity is another navigation technique that exploits signal communications already in existence, however similar to GPS-navigation SoOP requires an unobstructed view of transmission

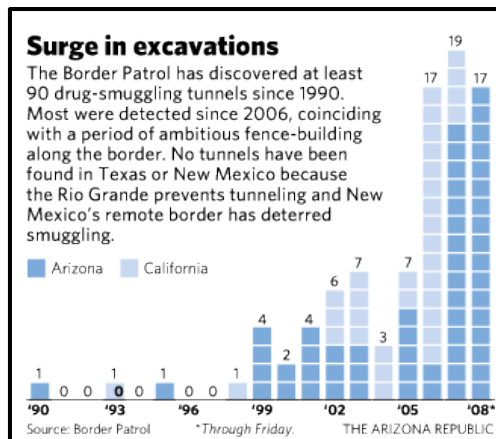


Figure 1.1: Surge in Tunnel Excavations [18]

sources which is not available underground. Dead reckoning based odometry is an appealing technique due to cost and simplicity but is not reliable due to error accumulations that would be magnified in the underground environment terrain [11]. Since the underground tunnel environment had not been previously characterized, navigation based on *a priori* maps is not possible. Laser-based navigation is a viable option that provides precision navigation solutions but simply due to the cost of the system is not viable for underground navigation.

Vision aided navigation on autonomous mobile vehicles dates back to the late 1970s, when Moravec and Gennery used feature tracking for visual correction [15]. Visual odometry has also been used successfully with autonomous aircraft and underwater vehicles [2] [26]. Despite the advances, use of visual techniques and quality research in the application area using robots has been slow. Optical flow has been well researched in theory but not in actual use and the practical applications of the theory have yet to be explored in detail [5].

A possible solution to navigating in GPS-denied environments is optical flow odometry based image registration, which can be used to constrain inertial sensor drift. Such an image-aided navigation system was developed by Lockheed Martin in the 1960s for military purposes and has continued to evolve to subpixel registra-

tion today [19] [35]. Although there have been significant advances in optical flow odometry there is a lack of research in the application to navigation.

## ***1.2 Scope***

This effort develops and demonstrates a navigation technique to augment stand-alone inertial sensors for improved navigation using optical flow odometry for improved tracking of system position and heading. This research will determine if optical flow based image registration technique is a viable solution for accurate pose navigation estimation. The scope of this thesis is limited to the development of a new registration algorithm design that registers images varying in translation only.

## ***1.3 Organization***

This thesis is organized in the following manner. A mathematical overview and background material in the area of robot navigation are presented in Chapter II. The methodology used to develop the optical flow based image registration technique will be covered in depth in Chapter III with detailed results and analysis in Chapter IV. Recommendations for future work and final remarks will conclude this thesis in Chapter V.

## II. Background

This chapter presents the mathematical and technical information necessary to fully develop an optical flow odometry navigation solution. The chapter begins by defining standard mathematical notation which is used throughout this document. Next, the navigation reference frames are defined and a mathematical technique for transforming coordinate reference frames will be defined. A basic understanding of inertial navigation will follow. Finally, a review of image registration techniques followed with current GPS denied navigation applications will be presented.

### 2.1 *Mathematical Background*

The mathematical notation used throughout this thesis will be presented in this section.

- Scalars: Scalars are represented with lower or upper case italics font. (e.g.,  $x$  or  $X$ )
- Vectors: Vectors are represented by lower case font with a vector bar over the variable. (e.g.,  $\bar{x}$ )
- Matrices: Matrices are denoted by upper case bold font. (e.g.,  $\mathbf{X}$ )
- Direction Cosine Matrices: Direction cosine matrices from frame  $a$  to frame  $b$  are denoted by  $\mathbf{C}_a^b$
- Reference Frame: If a vector is expressed in a specific reference frame, a superscript letter is used to designate the reference frame (e.g.,  $\bar{x}^n$  is the vector  $\bar{x}$  in the  $n$ -frame)

*2.1.1 Coordinate Systems and Transformations.* Navigation is an ancient skill or art which has become a complex science. It is essentially about finding the way from one place to another and there are a variety of means in which this can be achieved [10]. Navigation is performed in relation to a frame of reference (e.g., following directions on a map based on local surroundings). Navigation systems have



multiple applications but are often developed to enable a vehicle to correctly determine a navigation solution. Coordinate systems and transformation provide the foundation for defining the position and orientation of the navigation state relative to the applicable environment.

*2.1.1.1 Reference Frames.* Position can be described in terms of a variety of coordinate systems called reference frames, each from a different perspective. For instance, describing movement with respect to the earth requires a different frame of reference than movement on a Cartesian plane. The foundation to navigation is knowledge of position and orientation relative to a specific reference frame. Understanding reference frames allows for changes in position and orientation of a vehicle to be quantified. The following three-axis, right-handed reference frames are used in this research [38] [40].

- Inertial frame (*I*-frame): The true inertial frame is a theoretical reference frame defined by a non-accelerating, non-rotating, three-axis, right-handed coordinate system with no predefined origin or orientation where Newton's laws of motion apply.
- Earth-Centered frame (*i*-frame): The Earth-Centered frame is an orthonormal three-axis, right-handed coordinate system with origin at the center of mass of the Earth. The  $x$  and  $y$  axis are located on the equatorial plane, aligned with the fixed stars. The *i*-frame is a non-rotating frame but does accelerate with respect to the *I*-frame.
- Earth-Centered Earth-Fixed frame (*e*-frame): The Earth-Centered Earth-Fixed frame is a three-axis, right-handed coordinate system with the origin at the Earth's center of mass. The *e*-frame is rigidly attached to the Earth, with the  $z$ -axis aligned with the north pole,  $x$ -axis aligned with the intersection of the Greenwich meridian and the  $y$ -axis aligned with the equatorial plane pointing toward the 90 degrees east longitude. Figure (2.1) presents the *e*-frame.

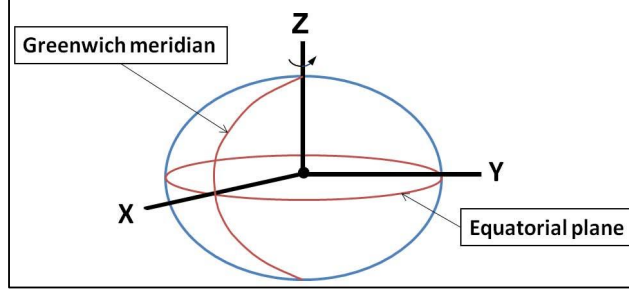


Figure 2.1: Earth-Centered Earth-Fixed Frame.

- Local-Level Navigation frame ( $n$ -frame): The Earth-fixed Local-Level Navigation frame is a three-axis, right-handed coordinate system with origin located at a point defined on the vehicle body. The vehicle's fixed navigation frame  $(x_n, y_n, z_n)$  point in the north, east, and down (NED) directions, respectively. The down direction is defined to align with the local gravity vector. Figure 2.2 presents the  $n$ -frame.
- Body frame ( $b$ -frame): The body frame is a three-axis, right-handed coordinate system rigidly attached to the vehicle body with origin co-located with the navigation frame. Looking at the body from an aerial view; the  $x_b$ ,  $y_b$ , and  $z_b$  axes point out front, right and bottom of the body, respectively.
- Sensor frame ( $s$ -frame): The sensor frame is a three-axis, right-handed coordinate system. The origin of an individual  $s$ -frame is defined relative to each individual sensor. Within this research each origin of an  $s$ -frame is co-located with the  $b$ -frame, but each sensor has a unique  $x$ ,  $y$ , and  $z$  orientation.

*2.1.2 Coordinate Transformations.* Coordinate transformations are used to resolve measurements from one reference frame into another. Describing the relationship between two reference frames is fundamental to the process of inertial navigation [38]. The two coordinate transformations pertinent to this research are direction cosine matrices (DCM) and Euler angles. A DCM is a  $3 \times 3$  matrix used to describe a vector in a different coordinate frame according to

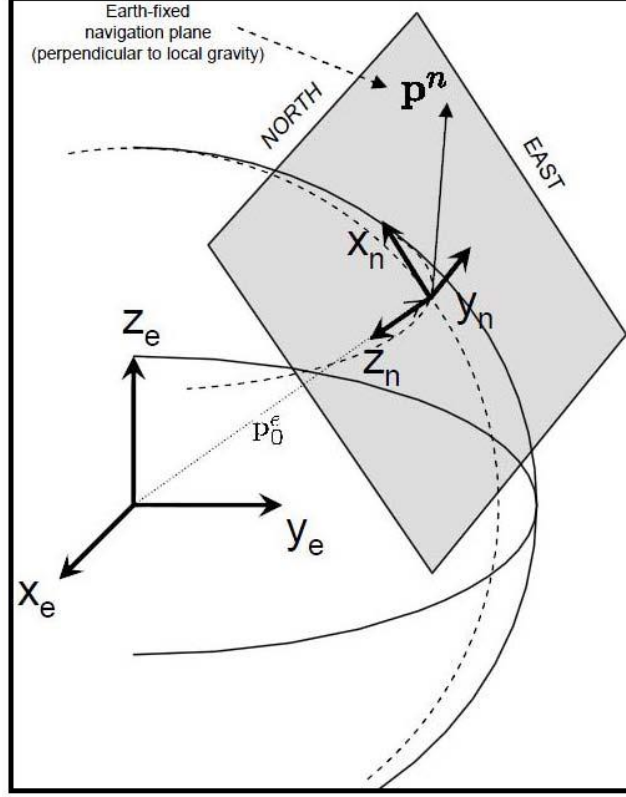


Figure 2.2: Earth-fixed Local-Level Navigation Frame.

$$\bar{r}^b = \mathbf{C}_a^b \bar{r}^a \quad (2.1)$$

where  $\bar{r}^a$  is a vector expressed in an arbitrary reference frame  $a$ .  $\bar{r}^b$  is the same vector described in the  $b$ -frame and  $\mathbf{C}_a^b$  is the DCM that describes three-dimensional rotation matrix between the  $a$ -frame and the  $b$ -frame. The element in the  $i$ -th row and the  $j$ -th column of  $\mathbf{C}_a^b$  represents the cosine of the angle between the  $i$ -th axis of the  $a$ -frame and the  $j$ -th axis of the  $b$ -frame [38].

Euler angles are angles that describe the relationship between two reference frames and are used to build the DCM. A series of three rotations about three different axes will transfer a system from one coordinate frame to another. Rotations of  $\psi$  about the  $z$ -axis,  $\theta$  about the  $y$ -axis and  $\phi$  about the  $x$ -axis can be described by three separate DCMs

$$\mathbf{C}_1 = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2)$$

$$\mathbf{C}_2 = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (2.3)$$

$$\mathbf{C}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (2.4)$$

When performing a transformation from the navigation frame ( $n$ -frame) to the body frame ( $b$ -frame) the axes may be expressed as the product of the three transforms as follows:

$$\mathbf{C}_n^b = \mathbf{C}_3 \mathbf{C}_2 \mathbf{C}_1 \quad (2.5)$$

Using the  $\mathbf{C}_n^b$  DCM, a vector  $\bar{p}^b$  in the body frame can be transformed to the navigation frame by

$$\bar{p}^b = \mathbf{C}_n^b \bar{p}^n \quad (2.6)$$

The inverse transformation from the body to the navigation frame is given by:

$$\mathbf{C}_b^n = \mathbf{C}_n^b{}^T = \mathbf{C}_1{}^T \mathbf{C}_2{}^T \mathbf{C}_3{}^T \quad (2.7)$$

## 2.2 Image Registration

Image registration is the process of taking two images of the same scene that are taken from different viewpoints or sensors and geometrically aligning them. Image registration can be divided into the following categories: correlation methods which use image pixel values directly [4], fast Fourier transform-based (FFT) methods which use the frequency domain [23], feature based methods which use low-level features such as edges and corners [11], and graph-theoretical methods which use high-level features such as specific objects or relationships between features [11].

The registration method used for this research utilizes the Fourier domain approach to match images that differ in translation. The Fourier method differs by conducting all computations in the frequency domain as opposed to the spacial domain [11] using the phase correlation technique [23]. Performing calculations in the frequency domain allows for excellent robustness against correlated and frequency-dependent noise.

*2.2.1 Digital Image Processing.* Digital image processing is the use of computer vision to process digital images. Modern digital technology makes DSP the cheapest and most versatile method of image processing. Digital image processing allows for a much wider range of algorithms than analog processing making it more efficient at simple tasks and the only practical solution for such tasks as feature extraction, pattern recognition, and image registration to name a few [24].

*2.2.2 Cross-Correlation.* One approach to image registration maximizes the correlation between two images by normalizing the cross-correlation, which is calculated directly in the spatial domain. Normalized cross-correlation is typically used in order to make the similarity measure independent of uniform changes in the image intensities. [41]

Given two images  $I_1$  and  $I_2$ , the normalized cross-correlation,  $c$ , of  $I_1$  and  $I_2$  at a relative shift  $(\Delta x, \Delta y)$  is

$$c = \frac{\sum_{x=1}^N \sum_{y=1}^M (I_1(x, y) - \mu_1)(I_2(x - \Delta x, y - \Delta y) - \mu_2)}{\sqrt{\sum_{x=1}^N \sum_{y=1}^M (I_1(x, y) - \mu_1)^2 (I_2(x - \Delta x, y - \Delta y) - \mu_2)^2}} \quad (2.8)$$

where  $\mu_1$  and  $\mu_2$  are the mean values (average pixel value over the entire image) of  $I_1$  and  $I_2$ , respectively. For image  $I_R$  that is  $m$  pixels tall and  $n$  pixels wide,  $\mu_R$  is calculated as

$$\mu_R = \frac{1}{mn} \sum_{i,j=1}^{n,m} I_R(i, j) \quad (2.9)$$

The cross-correlation is evaluated for each pixel combination between  $I_1$  and  $I_2$  and the 2D coordinate position of the maximum of the cross-correlation represents relative displacement between the images.

The cross-correlation technique in the spatial domain is computationally expensive due to the necessity of evaluating every displacement. The cross-correlation method is an adequate application for small amounts of data, but processing images using the cross-correlation method is unrealistic due to computational cost. The phase correlation method, the product of the Fourier transform of one image and the complex conjugate of the Fourier transform of the other image, drastically decreases the computational cost because the cross-correlation is done in the frequency domain instead of the spatial domain.

*2.2.3 Phase Correlation.* Phase Correlation Method (PCM) is a well known image registration method that was first proposed by Kuglin and Hines in 1975 [23]. PCM exploits the Fourier Shift Theorem property of Fourier Transform and estimates translational misalignment. This method is robust to frequency-dependent noise and dissimilar images, e.g., varying lighting or atmospheric conditions.

Consider two images  $I_1$  and  $I_2$  of the same scene that differ only by a displacement  $(\Delta m, \Delta n)$ . See Figure 2.3.

The corresponding Fourier transforms  $F_1$  and  $F_2$  are related by

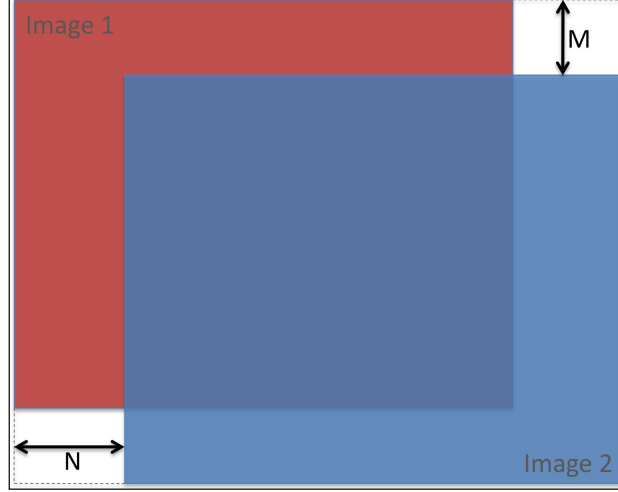


Figure 2.3: Image Shift:  $I_2$  is a Shifted Copy of  $I_1$ .

$$F_2(\omega_m, \omega_n) = F_1(\omega_m, \omega_n) e^{-j(\omega_m \Delta m + \omega_n \Delta n)} \quad (2.10)$$

Equation 2.10 shows that the two images have the same Fourier magnitude but a different phase difference directly corresponding to the displacement  $(\Delta m, \Delta n)$ .

Calculate the normalized phase correlation by taking the complex conjugate of the second image, multiplying the Fourier transforms together element wise, and normalizing the product element wise

$$R = \frac{F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)}{|F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)|} = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \quad (2.11)$$

Obtain the normalized cross-correlation by applying the inverse Fourier transform to  $R$

$$r = F^{-1}(R) \quad (2.12)$$

Determine the location of the shift between  $I_1$  and  $I_2$  by

$$(\Delta x, \Delta y) = \operatorname{argmax}_{(x,y)} \{r\} \quad (2.13)$$

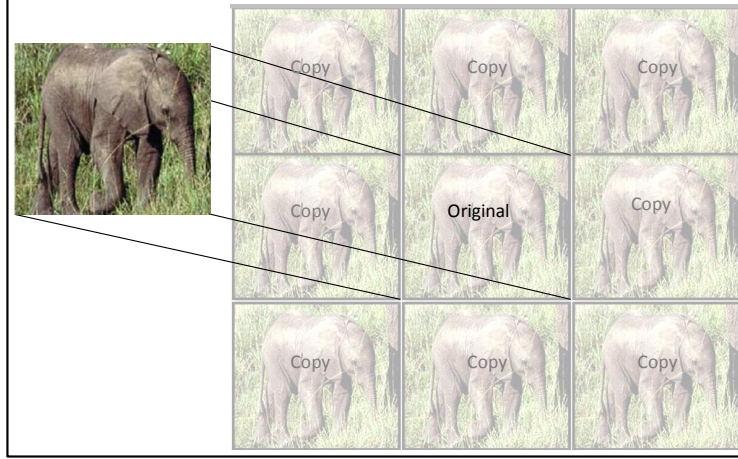


Figure 2.4: Periodic Extensions of the Original Image Due to Fourier Methods of Cross-Correlation.

The discrete convolution theorem presumes that the signal is periodic. Real data often consists of one non-periodic stretch of data with finite length [33]. Figure 2.4 illustrates the effects of the assumption that the signal, an image in this case, is periodic and shows the repetition of the image. This interference is called the wraparound effect [30].

To maintain the validity of the phase correlation method, it is necessary to find a workaround for the wraparound constraint. Since the convolution theorem assumes that the data is periodic and pollutes the image with extraneous data, a buffer zone needs to be added to the image to protect the data from these effects. Adding a buffer zone of zeros, called zero padding, will not introduce data that will interfere with the calculation of the displacement but will prevent the wraparound effect. Figure 2.5 illustrates the zero-padded original image.

The Phase Correlation method offers several remarkable properties: immunity to uniform variations of illumination, insensitivity to changes in spectral energy and, most importantly, excellent peak localization accuracy [39]. To establish a measure of correctness the curvature of the cross-correlation between the images is examined. The curvature points in the direction of maximum rate of change and allows the algorithm to differentiate between an actual shift and an erroneous shift. It is assumed that



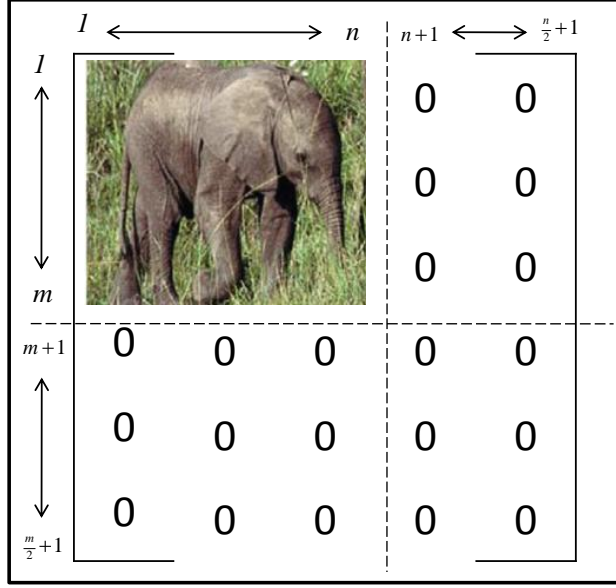


Figure 2.5: Zero-Padded Image To Prevent the Wraparound Effect.

the slope is small compared with unity so the approximation for the curvature of the cross-correlation,  $g(x, y)$ , can be found by taking the second derivative of the cross-correlation,  $r(x, y)$ , (Equation 2.14) [17].

$$g(x, y) = [\nabla^2(r(x, y))] \quad (2.14)$$

The use of the curvature overcomes the difficulty of being able to resolve the correct displacement and maintains the computational simplicity of the frequency based cross-correlation method.

Evaluation of the correctness of the cross-correlation peak can be assessed based on how much the curvature of the cross-correlation peak is above the noise floor using Equation 2.16. Dividing the maximum curvature value,  $\max(g(x, y))$ , by the mean of the curvature,  $\mu_A$ ,

$$\mu_A = \frac{1}{mn} \sum_{x,y=1}^{n,m} g(x, y) \quad (2.15)$$

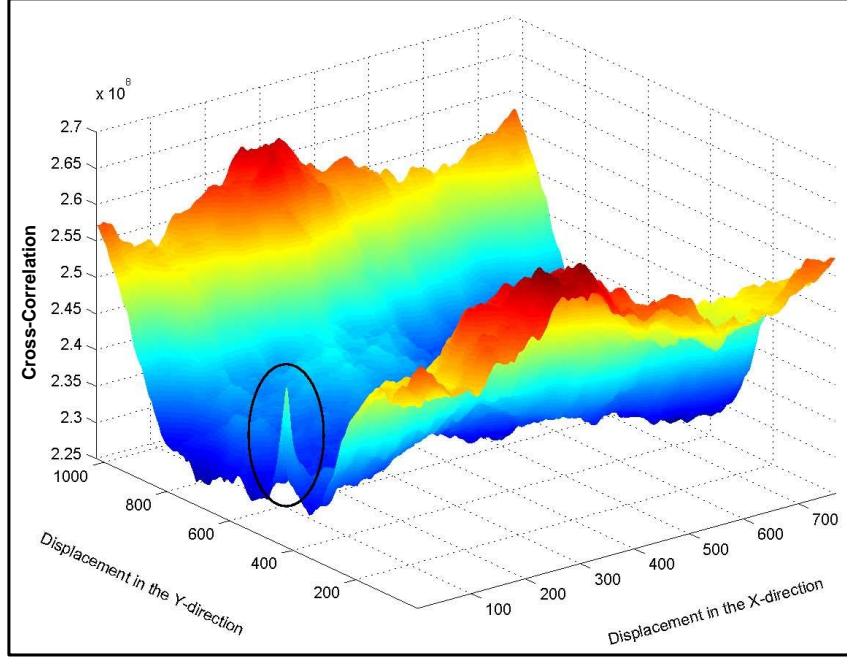


Figure 2.6: Cross-Correlation Plot Based on Fourier Methods: Difficult to Automatically Resolve Displacement Seen Highlighted in Black Circle.

provides a metric ( $A$ ) to assess the correctness of the displacement calculated by

$$A = \frac{\max(g(x, y))}{\mu_A} \quad (2.16)$$

It can be seen in Figure 2.6 that the displacement determined by the cross-correlation peak, highlighted in the black circle, of two images will not be resolved by evaluating the maximum value of the cross-correlation because it is not the maximum. The curvature of the cross-correlation as shown in Figure 2.7 resolves the displacement very clearly by evaluating the maximum rate of change of the cross-correlation.

### 2.3 Non-GPS Navigation Techniques

Over the past couple of decades the motivation to improve the ability to navigate in all environments has gained increasing importance. Precision navigation in GPS-denied environments is becoming a necessity in modern operations. GPS-navigation requires receiver line-of-sight, thus as missions continue to move underground, it is

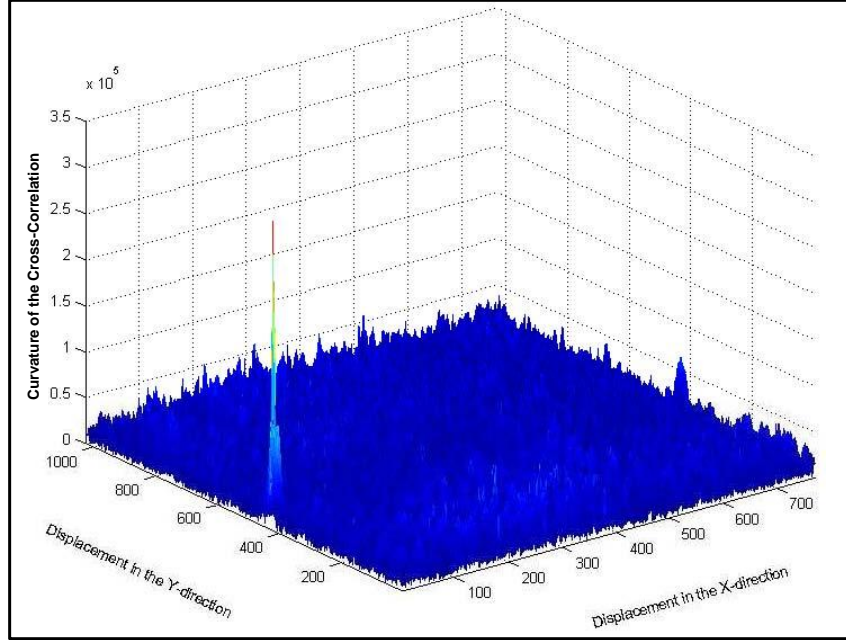


Figure 2.7: Curvature of the Cross-Correlation Plot: Displacement Easily Resolved With the Curvature of the Cross-Correlation.

not a viable solution. The problems in GPS availability necessitate the development of non-GPS dependent navigation techniques.

The advantages of non-GPS dependent navigation techniques increase the mobility and control by not limiting our reach to areas where line-of-sight to satellites is available, providing navigation redundancy where GPS is denied. It provides the capability to navigate and characterize unknown environments in areas where GPS cannot reach. Non-GPS navigation can be used to guide robotic platforms into regions of the battlefield where GPS is unavailable. Non-GPS navigation can provide U.S. and allied forces the ability to conduct timely reconnaissance, surveillance, and target acquisition in unfamiliar urban environments currently being investigated by the Dragon Runner project [37].

Recent literature has demonstrated numerous methods of precision navigation without GPS, from using signals of opportunity (SoOP) to vision aided approaches. The remaining sections outline several approaches for non-GPS navigation and highlight the advantages and disadvantages for underground robot navigation.

*2.3.1 Dead Reckoning.* A common method of calculating odometry is based on dead reckoning. Dead reckoning is the process of estimating current position based on a previously determined position. Initial mobile robotics used dead reckoning to estimate a robot's position and orientation with respect to a local reference frame placed in the environment. This is an attractive means of keeping track of position because it is inexpensive and can be done in real-time. Dead reckoning is usually accomplished with wheel, axle, or gear encoders by counting the number of rotations followed by converting the number of rotations into ground distance traveled. The primary short fall in of dead reckoning based upon odometry measurements is that each rotation does not necessarily correspond to ground distance traveled [9] [8]. The main short-coming of this approach is the skid-steering design of mobile robots. The skid-steering mobile robot is designed to have two differentially driving wheel pairs on each side of the robot but lacks a steering device, this design introduces uniform wheel movement that does not necessarily correspond to ground distance traveled [21]. Other forms of error that can be introduced with the dead reckoning method are gravel or muddy terrain when axle rotations do not directly relate to motion. Furthermore, the position estimations are based on the previously calculated positions, thus the errors of the process are cumulative and unbounded [38].

*2.3.2 Navigation Based on A Priori Maps.* Another navigation technique relies on *a priori* maps, a map that has predefined details about the environment. For this application, navigation is conducted by recognizing and matching features to localize current pose in the map [27]. This navigation technique is not a viable solution for underground environments as discussed in this thesis, because no previous knowledge of the environment is assumed.

*2.3.3 Laser-Based Navigation.* A wide range of laser-based sensors are currently being applied to the GPS-denied navigation problem [22] [27] [14] [6]. Laser sensors are able to capture fine angular and distance resolution, real time behavior with hundreds of point measurements per second and with low false positive rates.

Efficient algorithms exist for mapping and localization using this technique. However, image based sensors are often preferred due to laser technology costs with typical sensors costing an order or magnitude more [22].

*2.3.4 Signals of Opportunity Navigation.* Signals of opportunity based navigation estimates position using multi-lateration techniques similar to GPS. Pseudoranges are created using Time Difference of Arrival (TDOA) distances between a reference receiver and a mobile receiver, allowing the receiver to obtain a position estimate over time. Similar to the difficulties with using GPS, SoOP navigation requires the use of transmission sources such as Amplitude Modulation (AM) signal characteristics that are not viable in an underground setting [29].

*2.3.5 Visual Based Navigation.* Visual based navigation techniques are typically classified as either feature-based or optic flow-based [3] [16]. Feature-based methods determine the correspondence of features in a scene over multiple frames and optical flow-based techniques determine correspondence for a whole image between frames.

Feature tracking-based navigation methods have been proposed both for fixed-mount imaging sensors or gimbal-mounted detectors. Many feature tracking-based navigation methods exploit knowledge of the target location and solve the inverse trajectory projection problem [1]. Feature-based navigation techniques such as simultaneous localization and mapping (SLAM) determine changes in pose by matching key points between images. This technique is popular because it is robust to changes in scale, illumination, and rotation. The major downfall in implementing this technique in underground environments is the lack of features to detect and track [25].

Optic flow methods have been proposed generally for elementary motion detection, focusing on determining relative velocity, angular rates, or obstacle avoidance [20]. Optical-flow based odometry is useful for a variety of reasons, e.g., cameras are fairly compact and inexpensive and are able to be mounted on very small robots.

Also, with the rapid decreases in the cost and power requirement of computation, the expense of using visual odometry is often diminutive compared to other methods [36]. Even though there are existing visual odometry techniques based on geometric inference, which are highly accurate and work well in a variety of environments, this research will explore visual odometry methods which are based on determining pose changes without explicitly determining the scene or camera geometry. Due to the complexity and sensitivity of camera calibration with geometric visual odometry these techniques are often difficult to execute. Furthermore, the configuration parameters are specific to each camera and lens requiring modification for any changes made to the system.

## ***2.4 Summary***

The background chapter presented the mathematical and technical information necessary to fully develop an optical flow odometry navigation solution which will be used throughout this document. The navigation reference frames were then defined and a mathematical technique for transforming coordinate reference frames. Next, a review of image registration techniques followed with current GPS-denied navigation applications.

The methodology used to develop an optical flow based registration algorithm will be covered in depth in Chapter III with detailed results and analysis in Chapter IV. Recommendations for future work and final remarks will conclude this thesis in Chapter V.

### III. Methodology

The purpose of the optical flow based registration algorithm is to constrain drift present in an inertial sensor when GPS is not available. The development of a registration algorithm optimized for phase correlation between images as a vehicle moves through an underground environment is fundamentally an estimation problem involving mapping from an image scene at time ( $t$ ) to an image scene at time ( $t + 1$ ). This chapter outlines the optical flow algorithm used for pose estimation. The methods presented in this thesis for estimating vehicle pose using optical flow based image registration techniques follow the basic steps shown in Figure 3.1.

#### 3.1 Theory and Algorithms

*3.1.1 Phase One: Single Image Characterization.* Algorithm testing began with one large image, taking a series of known crops of the image and verifying that the correct shifts were calculated. To better explain this technique consider the following example. First, a 4111 x 6300 pixel image shown in Figure 3.2 is cropped into two 1024 x 768 pixel images shown in Figures 3.3 and 3.4. Figure 3.3 is the reference image, denoted by the solid rectangle in Figure 3.2, taken at time-step  $t$ . Figure 3.4 is considered the translated image taken at time-step  $t + 1$ , denoted by the dashed rectangle in Figure 3.2. In this example the difference between the reference image and the translated image is a shift of 0 pixels in the  $Y$ -direction and 200 pixels in the  $X$ -direction applied to the translated image.

Using image registration techniques discussed in Chapter II, the Fourier transform is taken of each cropped image and the phase correlation is assessed by multiplying the first cropped image by the complex conjugate of the second cropped image. The inverse Fourier transform is taken of the result and the maximum cross correlation is assessed resulting in the displacement between the two images. Figure 3.5 shows the cross-correlation between the two images, the peak represents the maximum correlation. The wraparound effect from the Fourier domain calculations can

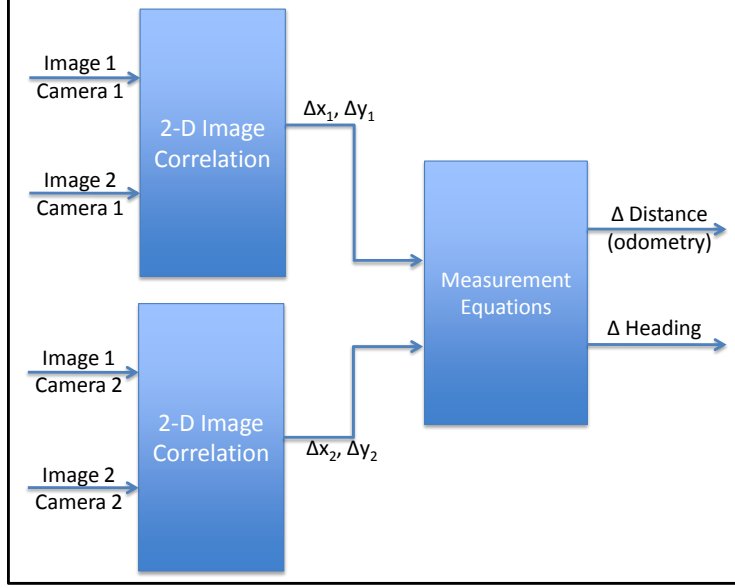


Figure 3.1: Optical Flow Based Registration Algorithm.

be seen by the peaks represented at a shift location of  $[200 \ 0]$  and  $[768 \ 0]$ . Effectively it is unknown if the actual shift was at  $[200 \ 0]$  or  $[200 - N \ 0]$ .

Curvature of the cross-correlation provides a measure to assess the correctness of the predicted shift by identifying how much the maximum rate of change of cross-correlation is above the average curvature of the cross-correlation. The displacement between image  $a$  and image  $b$  in Figure 3.6 can easily be inferred by using the red highlighted circles. The cross-correlation between the images, as shown in Figure 3.7, shows the displacement peak, highlighted by the black circle, being overshadowed by noise. The true displacement cannot be resolved automatically from the maximum of the cross-correlation. From Figure 3.7 it can be seen that there are multiple peaks that have a greater value than the actual peak, but the actual displacement peak is characterized by a sharp peak. Assessing the curvature of the cross-correlation plot will resolve which peak has the sharpest change in slope. Figure 3.8 resolves the noise and determines the actual displacement seen by the single peak. This method can also be used to determine if a displacement cannot be determined between two images. Figure 3.9 shows a cross-correlation plot where no sharp peaks are present,



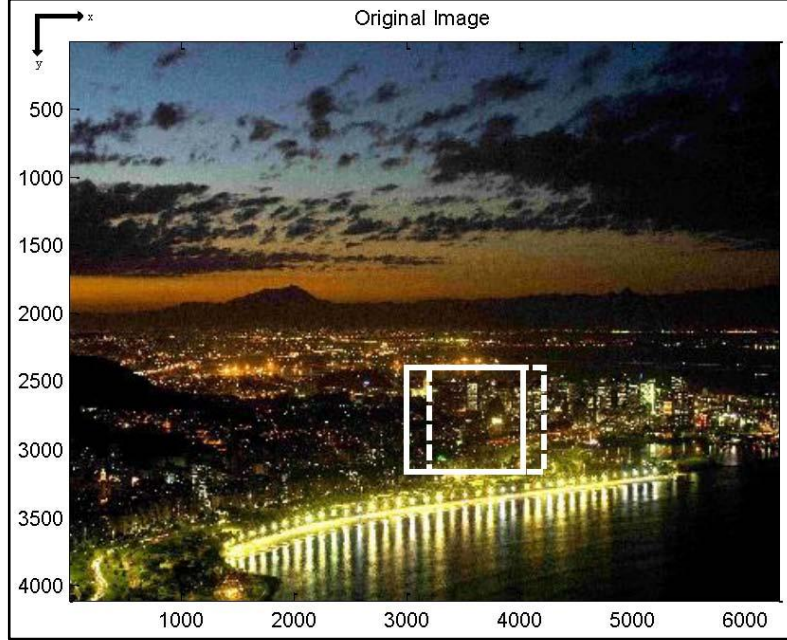


Figure 3.2: Single Large Image.

indicating that an accurate displacement cannot be determined. The curvature of the cross-correlation plot, as shown in Figure 3.9, indicates that the maximum value found from the cross-correlation is polluted with noise thus indicating a correlation between these images could not be determined. Automatic evaluation of the maximum of the curvature of the cross-correlation to determine the correctness of the prediction can be assessed using Equation 2.15, where  $A$  is a metric to determine the validity of the predicted displacement.

*3.1.2 Phase Two: Single Camera Characterization.* The next phase of the optical flow odometry based image registration algorithm characterization uses a single camera. Evaluating the algorithm with real-world images introduces images with pixel disparity at matching points. Pixel disparity is a result of slight changes in how the real world was mapped to each pixel between the images. This effect may not be noticeable to the naked eye, but due to the precision of each pixel, slight differences are determinable. The overcoming of these slight changes increases the robustness of the algorithm making it more reliable for real-world implementation.

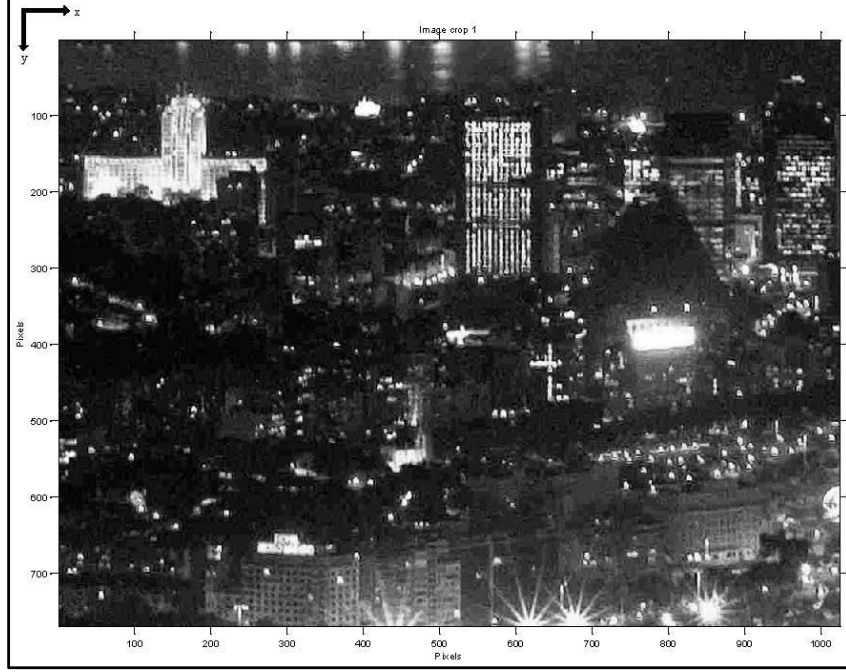


Figure 3.3: Image Crop 1.

A series of images are captured, as shown in Figure 3.11, and similar to the single image setup, two sequential images are evaluated in the frequency domain where the maximum of the cross-correlation resolves the displacement between the two images. Figure 3.12 shows image captured at time  $(t)$ , time  $(t + 1)$ , and time  $(t + 2)$ . By inspection of Figure 3.12, it is clear that the shift is in the positive  $Y$ -direction (assuming a orthogonal right-handed orthogonal reference frame).

*3.1.3 Phase Correlation.* Converting two sequential images from the spatial domain to the frequency domain where the phase shift is related to the relative translation in the spatial domain. The inverse Fourier transform of the normalized phase correlation defined by

$$R = \frac{F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)}{|F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)|} = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \quad (3.1)$$

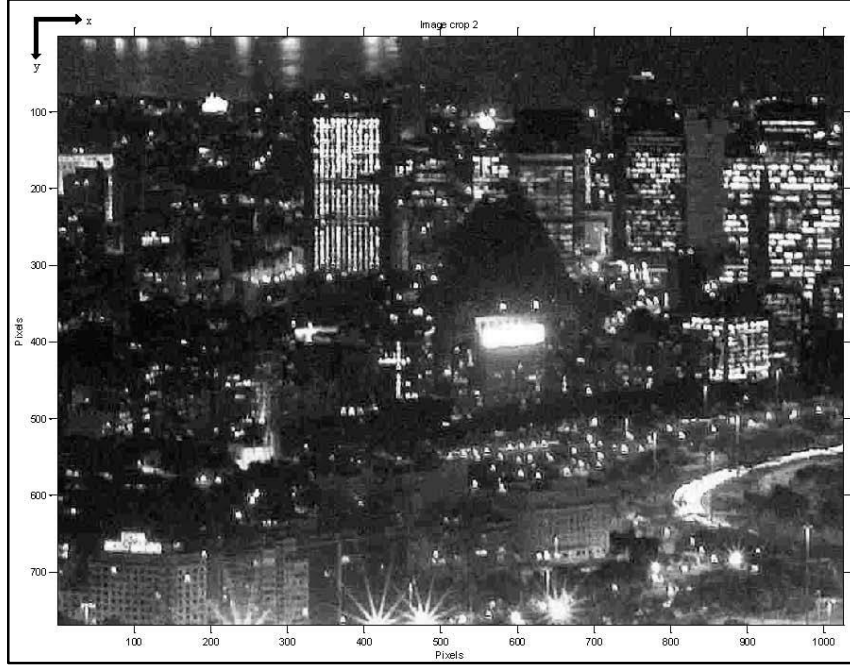


Figure 3.4: Image Crop 2.

resolves a maximum peak at the location of the displacement between the two images. Figure 3.13 illustrates the cross-correlation and estimates the displacement by the overall maximum peak.

*3.1.4 Phase Three: Two Camera Characterization.* From the readings of two cameras it is possible to compute the pose of a mobile vehicle independently from the kinematics. A description of how to estimate the vehicle pose considering that the cameras are in a linear orientation, separated by a distance  $D$ , as shown in Figure 3.14.

Each camera is able to measure a 2-dimensional translation in the coordinate system of the vehicle to which they are rigidly attached using the methods described in Section 3.1.2. If the vehicle makes an arc of circumference, it can be shown that each camera makes an arc of circumference, which is characterized by the same angle but different radius. An ambiguity may arise if only a single camera were used for odometry and heading. For example, Figure 3.15 shows two different paths that

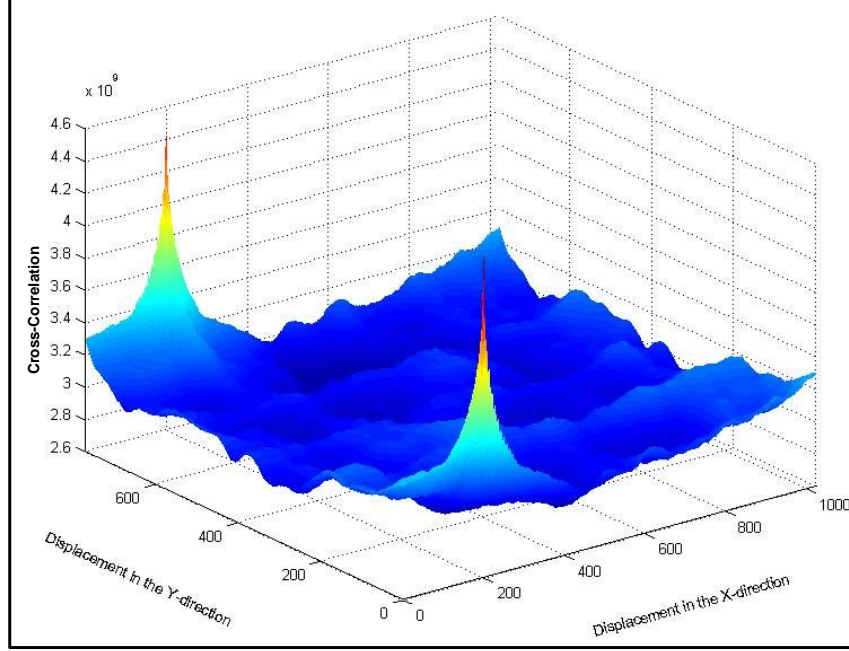


Figure 3.5: Single Image Cross-Correlation.

resolve identical  $\Delta x$  and  $\Delta y$  measurements. The path length,  $l$ , and the angle between the  $x$ -axis and the velocity vector,  $\alpha$ , is the same for both cases. Figure 3.15 (a) moved in a straight line, whereas Figure 3.15 (b) moved in an arc [7] [28].

In general, the following relationships can be made between the camera displacements and the path:

$$\Delta x = l \cos(\alpha) \quad (3.2)$$

$$\Delta y = l \sin(\alpha). \quad (3.3)$$

The angle  $\alpha$  between the  $x$ -axis of camera and the tangent is

$$\alpha = \arctan\left(\frac{\Delta y}{\Delta x}\right). \quad (3.4)$$

The length of the arc,  $l$ , can then be resolved using

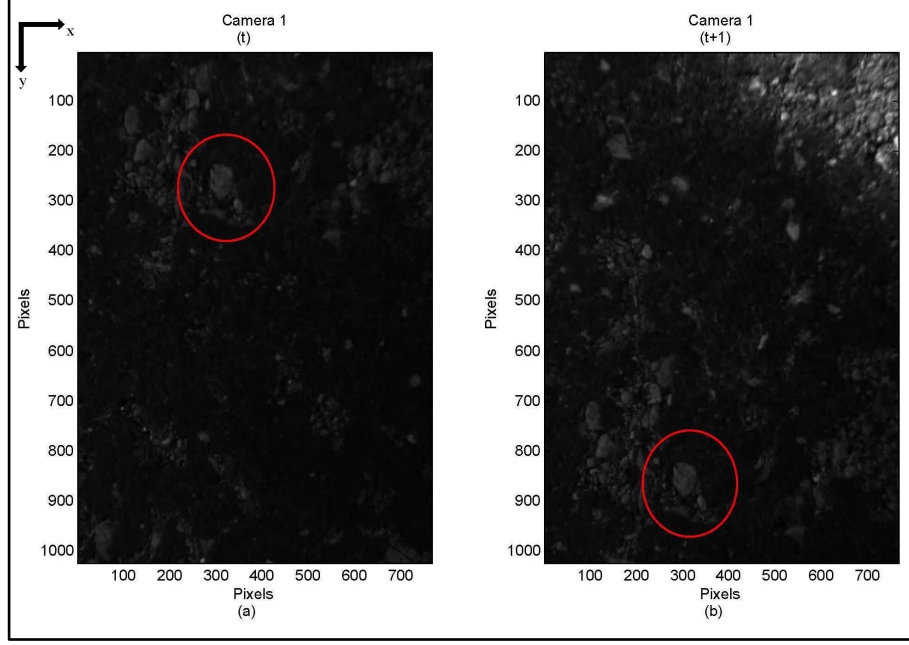


Figure 3.6: Determination of Shift Example.

$$l = \begin{cases} \Delta x & \text{for } \alpha = 0, \pi \\ \frac{\Delta y}{l \sin(\alpha)} & \text{otherwise.} \end{cases} \quad (3.5)$$

Over a short sampling period, it is assumed that the vehicle moves with constant tangential and rotational speeds, therefore during a sampling period the movement can be approximated by the arc of circumference. The arc of circumference can be described by  $x$  and  $y$  coordinates as well as the rotational angle  $\Delta\theta$ . Each camera resolves a  $\Delta x$  and  $\Delta y$  and the distance between the cameras is held constant at  $D$ . This means that the line that joins the cameras should always produce the same displacement value. From these readings it is possible to calculate the vehicle pose changes in terms of  $\Delta x$ ,  $\Delta y$ , and  $\Delta\theta$ . The cosine rule is applied to the triangle made by joining the radius of each camera ( $r_1$  and  $r_2$ ), the joining line ( $D$ ), and the angle between them ( $\gamma$ ) as shown in Figure 3.16. While  $\gamma$ , is described by the radii of both cameras it can be computed by  $\gamma = |\alpha_1 - \alpha_2|$ .

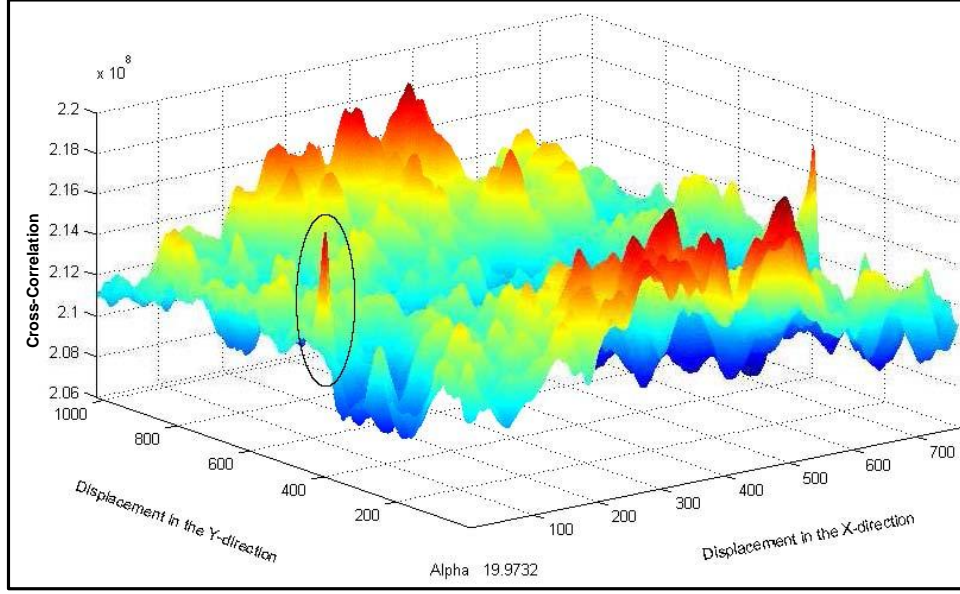


Figure 3.7: Cross-Correlation: Actual Shift is Overshadowed.

$$D^2 = r_1^2 + r_2^2 - r_1 r_2 \cos(\gamma) \quad (3.6)$$

The ratio between the arc length ( $l$ ) and the arc angle ( $\theta$ ) describe the radius ( $r$ ) of an arc by Equation  $l = r\theta$ . The two cameras are associated to arcs under the same angle, which directly corresponds to a change in orientation of the vehicle. The radii are described by

$$r_1 = \frac{l_1}{|\Delta\theta|}, \quad (3.7)$$

$$r_2 = \frac{l_2}{|\Delta\theta|}. \quad (3.8)$$

Substituting Equations (3.7) and (3.8) into Equation (3.6) results in an expression for change in orientation ( $\Delta\theta$ ) described by

$$\Delta\theta = \frac{\sqrt{l_1^2 + l_2^2 - 2 \cos(\gamma) l_1 l_2}}{D} \cdot \text{sign}(\bar{y}_2 - \bar{y}_1). \quad (3.9)$$

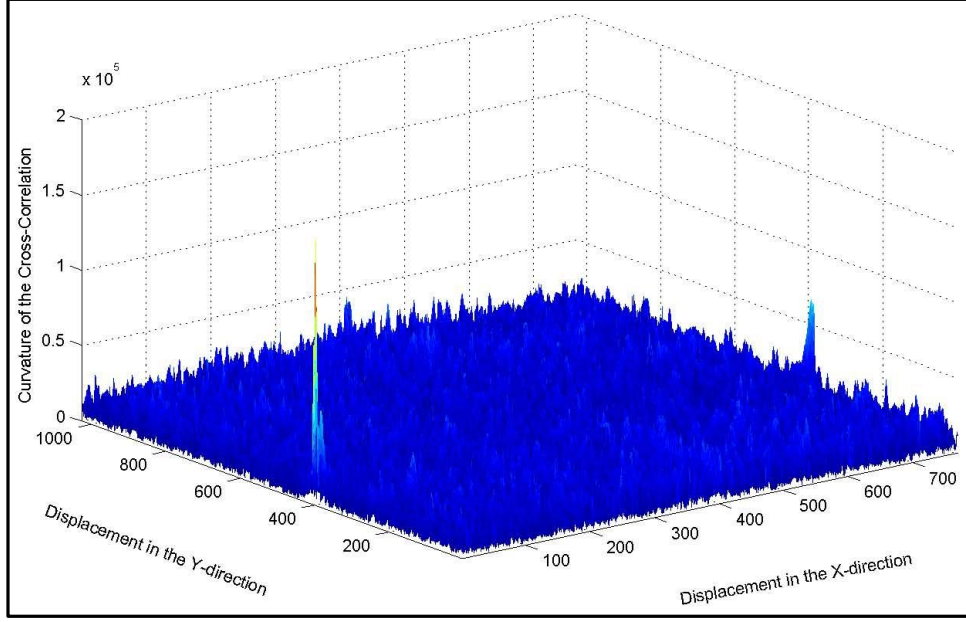


Figure 3.8: Curvature of the Cross-Correlation.

The vehicle coordinate frame, the  $b$ -frame, is described in Figure 3.14; it is an orthogonal frame with the two cameras rigidly attached to the vehicle body. The cameras are in a linear orientation separated by a distance  $D$ . The formulas for calculating the coordinates at the end of the sampling period are

$$x'_1 = r_1(\sin(\alpha_1 + \Delta\theta) - \sin(\alpha_1)) \cdot \text{sign}(\Delta\theta), \quad (3.10)$$

$$y'_1 = r_1(\cos(\alpha_1) - \cos(\alpha_1 + \Delta\theta)) \cdot \text{sign}(\Delta\theta) - \frac{D}{2}, \quad (3.11)$$

$$x'_2 = r_2(\sin(\alpha_2 + \Delta\theta) - \sin(\alpha_2)) \cdot \text{sign}(\Delta\theta), \quad (3.12)$$

$$y'_2 = r_2(\cos(\alpha_2) - \cos(\alpha_2 + \Delta\theta)) \cdot \text{sign}(\Delta\theta) + \frac{D}{2}. \quad (3.13)$$



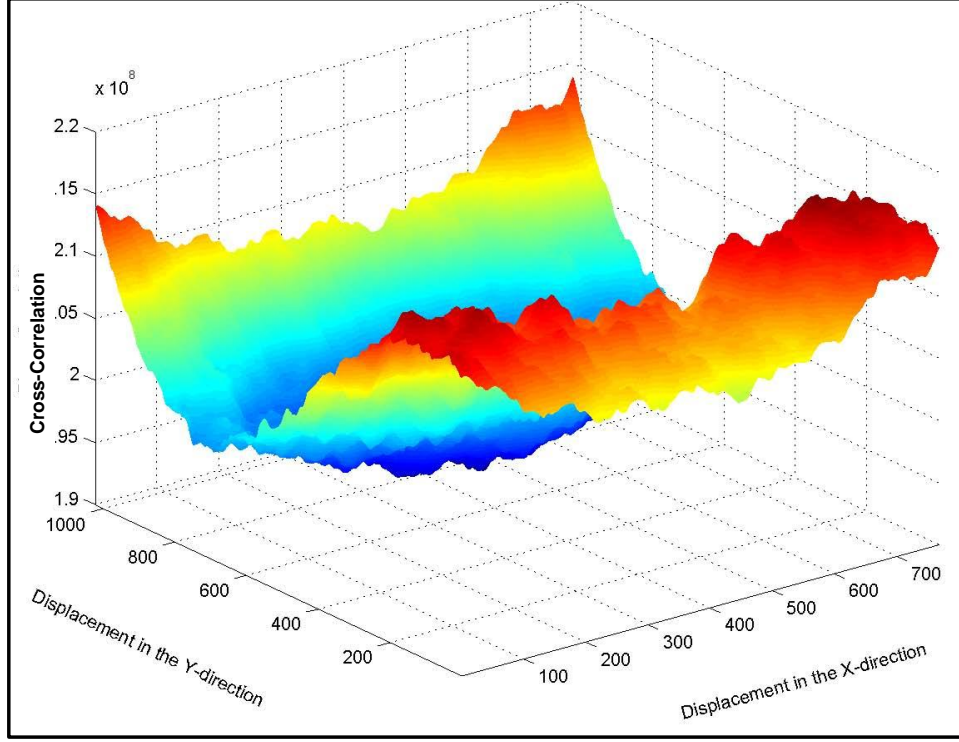


Figure 3.9: Cross-Correlation: No Correlation Between Images Determined.

The change in vehicle position in the  $x$  and  $y$  direction can be resolved using

$$\Delta x = \frac{x'_1 + x'_2}{2}, \quad (3.14)$$

$$\Delta y = \frac{y'_1 + y'_2}{2}. \quad (3.15)$$

The position of the vehicle, in the  $b$ -frame, at time  $t + 1$  can be computed by knowing the coordinates at time  $t$  and the relative movement carried out during the period from  $t$  to  $t + 1$  with

$$X_{t+1} = X_t + \sqrt{\Delta x^2 + \Delta y^2} \cos \left( \Theta_t + \arctan \left( \frac{\Delta y}{\Delta x} \right) \right), \quad (3.16)$$

$$Y_{t+1} = Y_t + \sqrt{\Delta x^2 + \Delta y^2} \sin \left( \Theta_t + \arctan \left( \frac{\Delta y}{\Delta x} \right) \right), \quad (3.17)$$



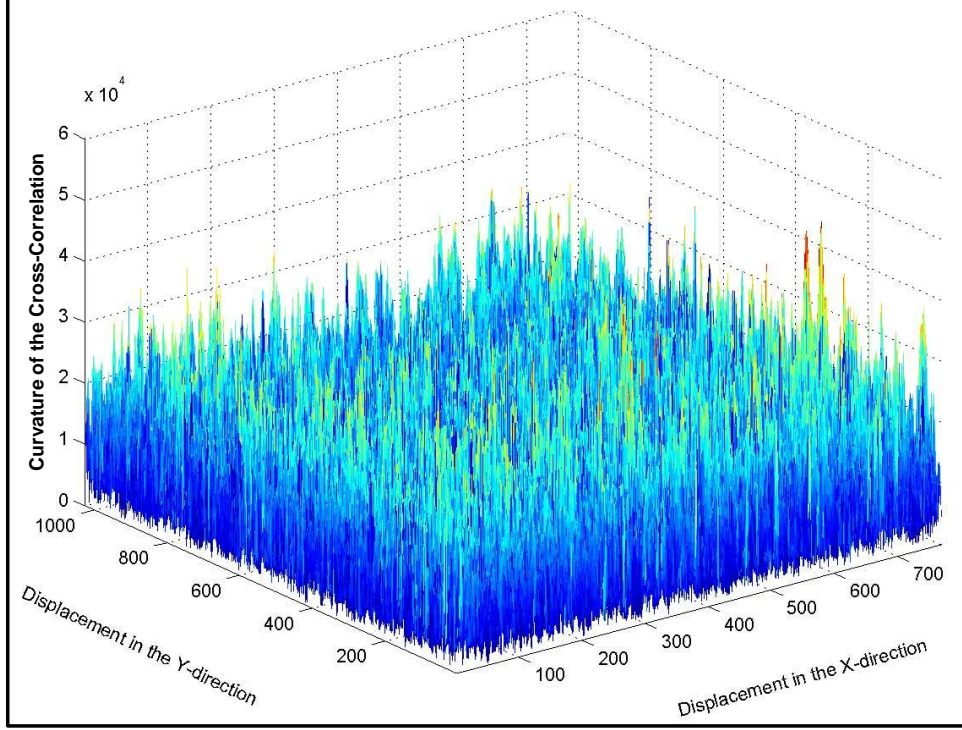


Figure 3.10: Curvature of the Cross-Correlation: No Correlation Between Images.

$$\Theta_{t+1} = \Theta_t + \Delta\theta \quad (3.18)$$

This process describes how the relative pose of the vehicle is resolved by applying the displacement outputs from each camera cross-correlation to the algorithm described in Section 3.1.4. No previous research has fused the technique of image registration from two downward looking cameras applied as an odometric sensor. The hardware setup and results to test the theory of this algorithm will be presented in Chapter IV.

### 3.2 Performance Metrics

The primary metric to evaluate the performance of the optical flow based registration algorithm is the evaluation of the absolute value of difference between the predicted  $(\Delta X_p, \Delta Y_p)$  and the truth  $(\Delta X_t, \Delta Y_t)$



Figure 3.11: Optical Flow From Single Camera Image Processing.

$$\text{X-error} = |\Delta X_p - \Delta X_t|, \quad (3.19)$$

$$\text{Y-error} = |\Delta Y_p - \Delta Y_t|. \quad (3.20)$$

To represent the multidimensional displacement, the Euclidean distance ( $l_2$  Norm) between the predicted  $(\Delta X_p, \Delta Y_p)$  and the truth  $(\Delta X_t, \Delta Y_t)$

$$\text{Error} = \sqrt{(\Delta X_p - \Delta X_t)^2 + (\Delta Y_p - \Delta Y_t)^2}, \quad (3.21)$$

also provides another metric for evaluation.

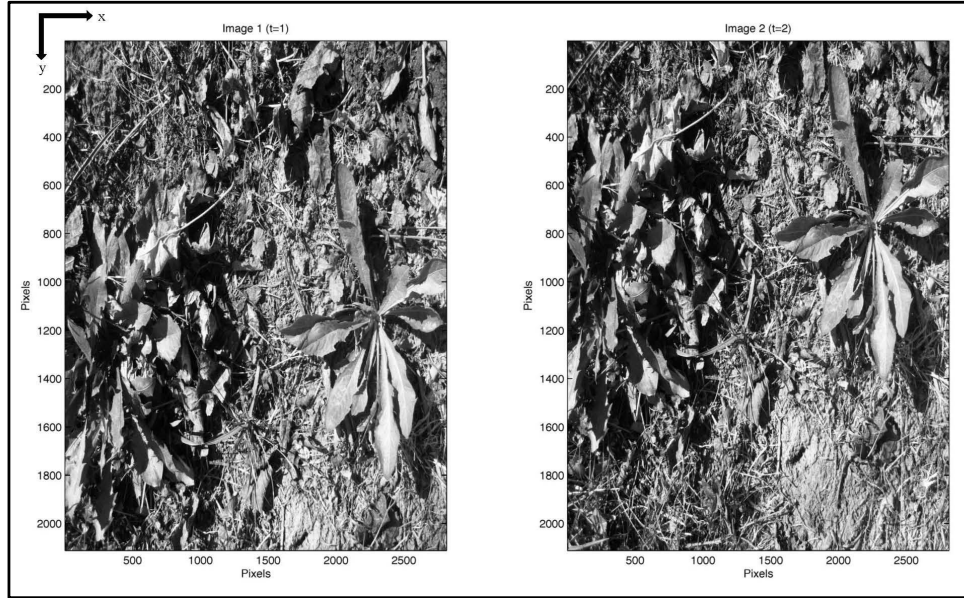


Figure 3.12: Image Captures From Single Camera Image Processing.

### 3.3 Summary

Chapter III presented the methodology used to develop an optical flow based registration algorithm. Chapter IV presents the hardware and testing parameters used to test the optical flow based registration algorithm followed by detailed results and analysis. Finally, recommendations for future work and final remarks will conclude this thesis in Chapter V.

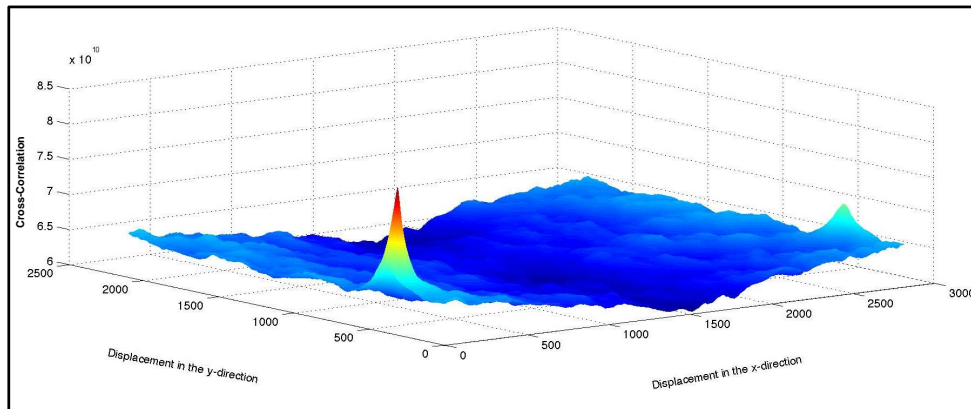


Figure 3.13: Cross-Correlation from Single Camera Image Processing.

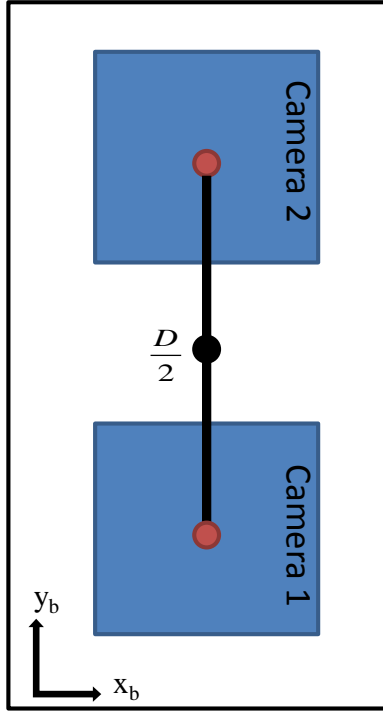


Figure 3.14: Positioning of the Cameras in the Body Frame. Camera 1 and Camera 2 are Parallel to Each other and Separated by a Distance  $D$ .

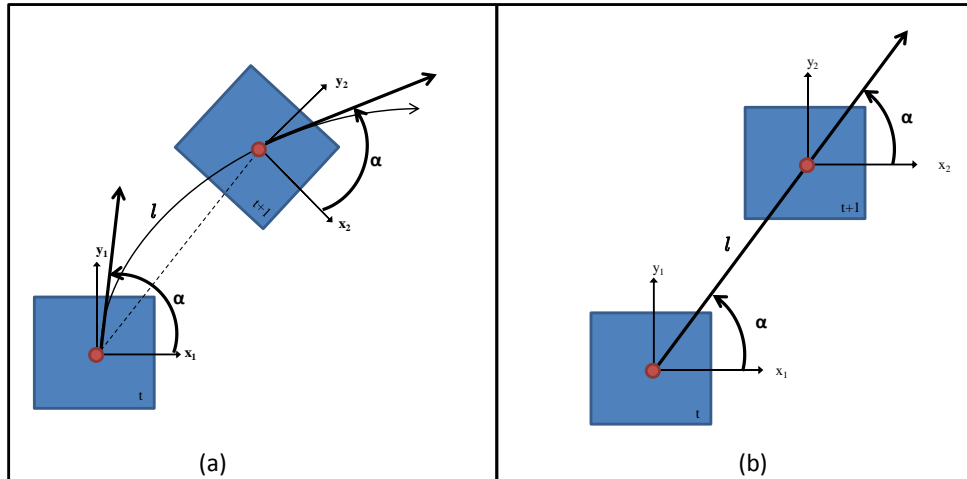


Figure 3.15: Two different Paths That Return the Same Readings.

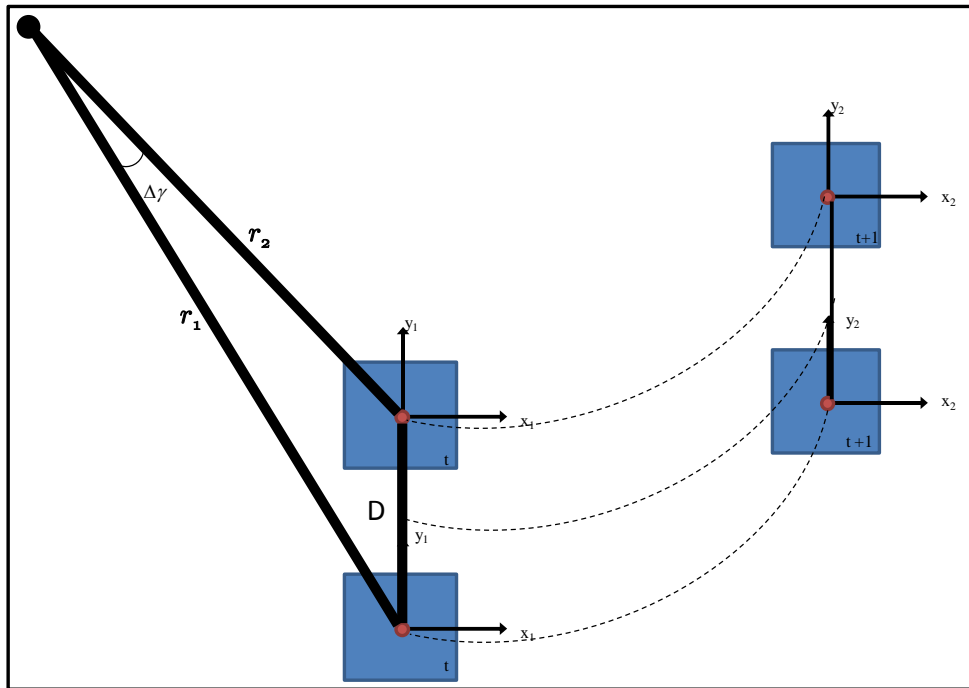


Figure 3.16: Triangle Used to Resolve  $\Delta\theta$  Based on the Cosine Rule.

## IV. Analysis and Results

This chapter presents hardware and testing parameters used to test the optical flow based registration algorithm. An analysis of simulated as well as real-world results are then presented, followed by an analysis of the results.

### 4.1 *Hardware Overview*

The navigation system is shown in Figure 4.1, the navigation reference system is highlighted in green and the image-aided navigation system is highlighted in purple. The navigation reference system consisted of a tactical grade Inertial Measurement Unit (IMU), GPS receiver and a Synchronous Position, Attitude and Navigation (SPAN) receiver. The image-aided navigation system used consisted of a navigation computer and a pair of digital cameras.

*4.1.1 Digital Cameras.* The digital cameras are both PixeLINK PL-A741 machine vision cameras. The PL-A741 camera uses a monochrome complementary metal oxide semiconductor (CMOS) imaging sensor with  $1280 \times 1024$  pixel (1.3 mega pixel) resolution. The pixels are  $6.7 \mu\text{m}$  square which results in a sensor size of  $8.576 \text{ mm} \times 6.921 \text{ mm}$ . The cameras are paired with VS Technology Corp SV-0514MP lenses. These lenses feature high resolution with low distortion and lock screws for focus and iris with a maximum sensor size of  $2/3$  inch [31]. The camera communicates with a laptop computer using an IEEE-1394a (FireWire) interface, which collect images at 10 Hz. Finally, the camera includes a global shutter which exposes all of the pixels simultaneously. The complete PL-A741 specifications are listed in [32]. The two cameras are mounted on a rigid bracket pointed downward (toward the ground) at 90 degrees relative to the vehicle, on the right side, 14.29 cm from the ground. They were spaced 45.4 centimeters apart. Figure 4.2 shows a diagram of the golf cart setup.

*4.1.2 Tactical Grade IMU.* The Honeywell HG1700 is a tactical-grade, strapdown IMU. The unit consists of a GG1308 ring-laser gyroscope and triad of RBA-500 accelerometers which produce measurements at 100 Hz. The gyroscopes

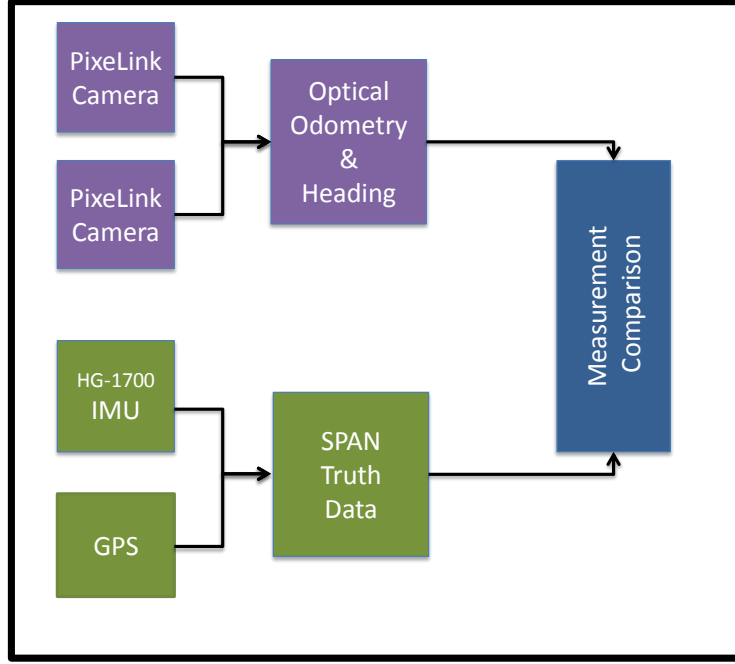


Figure 4.1: Implementation Block Diagram.

and accelerometers have a dynamic range of  $\pm 1000$  degrees per second and  $\pm 50g$ , respectively.

*4.1.3 SPAN.* Novatel, Inc's synchronized position, attitude and navigation (SPAN) receiver was used to interface GPS and the tactical grade IMU. The SPAN provided the truth source for evaluation of the optical flow based odometry algorithm.

*4.1.4 Vehicle Platform.* The coordinate frame for the system is shown in Figure 4.2. The system is configured in a right-hand orthogonal manner with vehicle movement in the forward direction being positive  $y$  relative to the  $b$ -frame. The origin of the  $b$ -frame is designed to coincide with the origin of the SPAN  $s$ -frame. Also, the origin of the SPAN is rotated to align with the  $b$ -frame. The  $x$ ,  $y$ , and  $z$  coordinates coincide with the right hand rule. Observe that when the vehicle is moving in the forward direction only, the positive  $y$  direction, it results in a positive  $m$  pixel movement and a right hand turn would result in a positive  $x$  measurement and a positive  $n$  pixel movement.

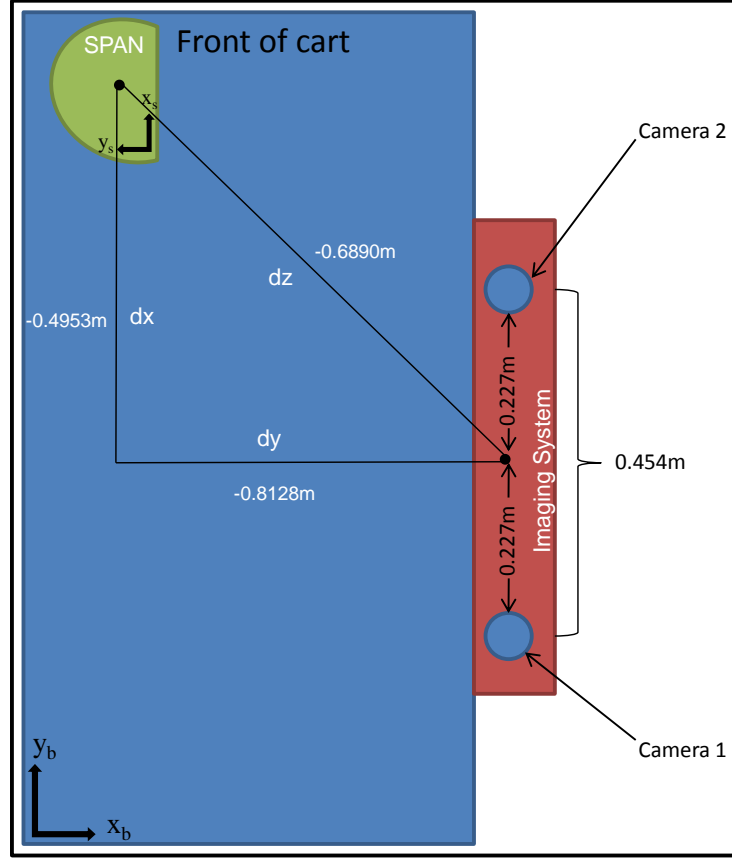


Figure 4.2: Vehicle platform setup.

Note that the imaging system and the SPAN are both mounted on a rigid body, each experiencing the same roll ( $\phi$ ), pitch ( $\theta$ ), and yaw ( $\psi$ ) as the vehicle. The imaging system is translated from the SPAN by the vector  $\mathbf{p}_{camera\ bar}^b = (-0.4853, -0.8128, -0.6890)$  meters.

## 4.2 Simulation Results

The outdoor environment was chosen to test the optical flow based algorithm implemented with two camera collection. The collect was taken during the day and covered the path seen in Figure 4.3. Images were captured at a rate of 10 Hz during the outdoor collect.



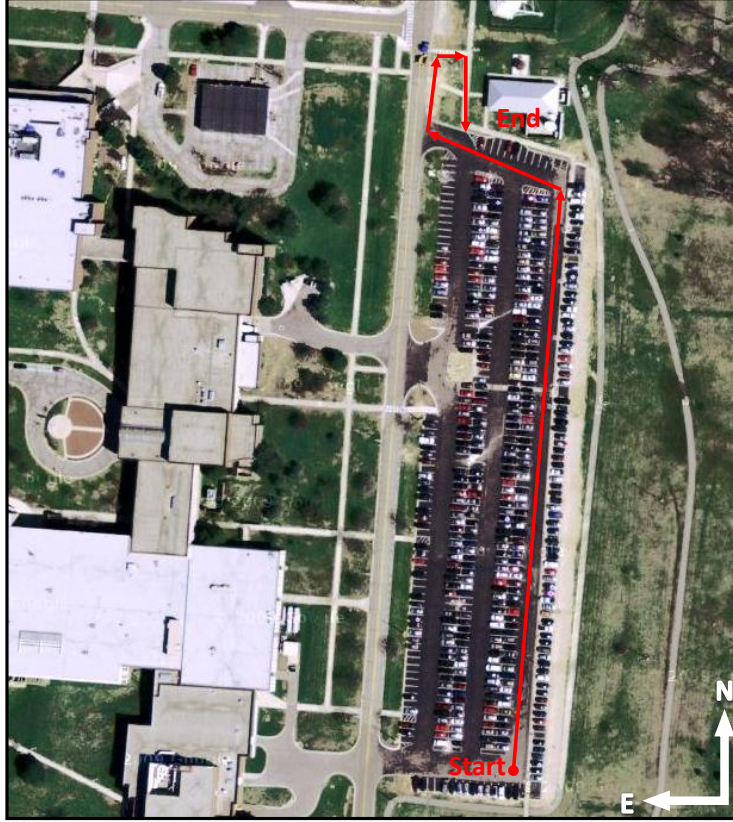


Figure 4.3: Vehicle platform setup.

The post-processing and optical flow algorithm is completed using **Matlab**<sup>®</sup> R2010a and follows the steps outlined in Chapter III. To evaluate the performance of this implementation an alpha ( $A$ ) parameter sweep was performed to determine the optimal threshold for accurate displacement detection, the predicted path was compared to the truth path and evaluated with the accuracy statistics.

*4.2.1 Peak Parameter Sweep.* Recall from Chapter III using image registration techniques and evaluating the maximum of the cross-correlation estimates the displacement between two images. Curvature of the cross-correlation provides a measure to assess the correctness of the calculated shift. Evaluation of the correctness of the cross-correlation peak can be done based on how much the maximum calculated cross-correlation value is above the average cross-correlation values, this parameter,  $A$ , is calculated using Equation 2.15.

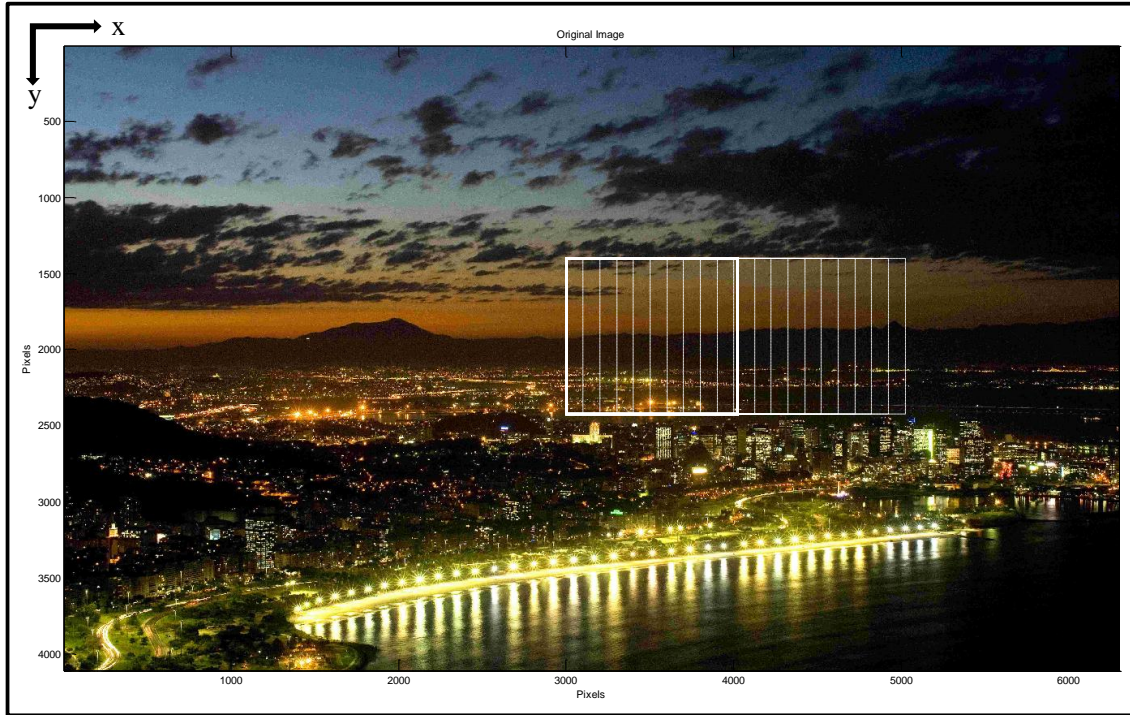


Figure 4.4: Alpha ( $A$ ) Parameter Sweep For X-Direction Test Setup.

The Peak Parameter Sweep used the single image test setup, previously discussed in Section 3.1.1, using one large image (4111 x 6300 pixels) shown in Figure 3.2 and cropping two 1024 x 768 pixel images, denoted by the solid rectangle and the dashed rectangle in Figure 3.2. The alpha parameter ( $A$ ) was evaluated at 1000 single pixel shifts in the positive  $x$ -direction (Figure 4.4), positive  $y$ -direction (Figure 4.5), and diagonal positive  $x$ -direction and  $y$ -direction (Figure 4.6). The alpha parameter is assessed in the positive direction only because it is assumed that the vehicle will always be moving in a positive  $y$ -direction or not moving. Figure 4.7 shows the average alpha value per pixel shift for the test runs. The red line represents a correct displacement prediction, the blue line represents an incorrect displacement prediction, and the black dot is the  $\alpha$  threshold value.

A threshold of  $A = 10$  is set to measure the correctness of the cross-correlation. The threshold value will allow the algorithm to ignore incorrect predicted shifts, thus producing a more accurate prediction vehicle pose.

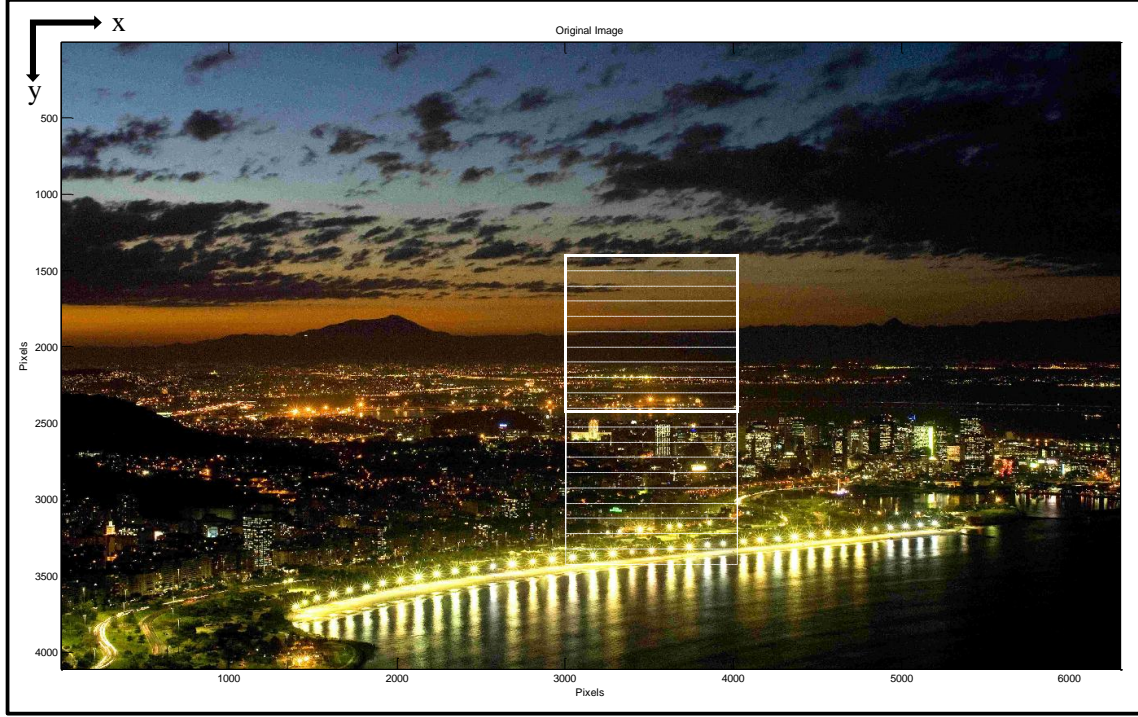


Figure 4.5: Alpha (A) Parameter Sweep For Y-Direction Test Setup.

*4.2.2 Displacement Error Calculation.* The primary metric to evaluate the performance of the optical flow based registration algorithm is the evaluation of the absolute value of the difference between the predicted ( $\Delta X_p, \Delta Y_p$ ) and the truth ( $\Delta X_t, \Delta Y_t$ ). This performance metric is calculated with Equations 3.19 and 3.20 for the  $X$ -direction and  $Y$ -direction respectively. The difference between the displacements in the  $X$ -direction are shown in Figure 4.8 and Figure 4.9 for the  $Y$ -direction. Tables 4.1 and 4.2 present the statistics for each direction of error distributions respectively.

Table 4.1: Statistics for Difference Between Predicted X-Displacements and Truth X-Displacements.

Parameter	Value
Mean	3.099 cm
Std Dev	6.733 cm

Observe that the difference in both the  $X$ -direction and the  $Y$ -direction hover around the means respectively, with a few outliers. The outlier values can be explained



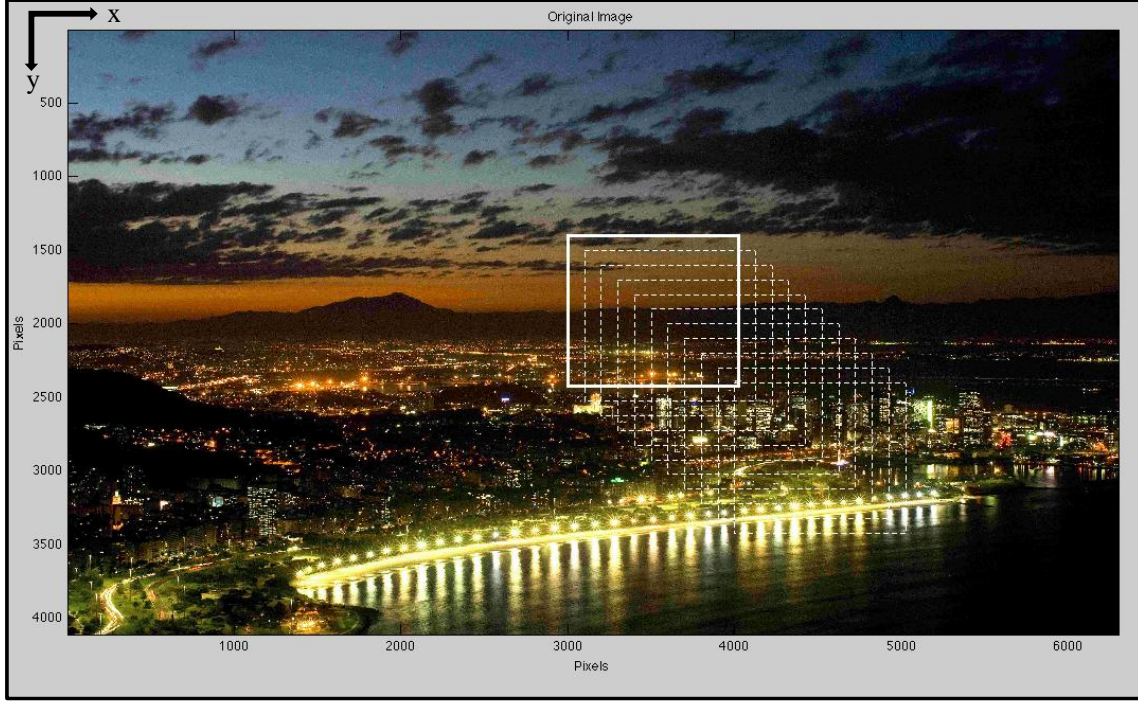


Figure 4.6: Alpha (A) Parameter Sweep For Diagonal Test Setup.

Table 4.2: Statistics for Difference Between Predicted Y-Displacements and Truth Y-Displacements.

Parameter	Value
Mean	3.454 cm
Std Dev	9.537 cm

by examining the average velocity of the run, shown in Figure 4.10. The velocity is steadily increasing and stabilizes around  $0.4 \frac{m}{s}$ , this translates to a consistent image overlap, around time step 350 there is an abrupt spike in the velocity causing the image overlap to be significantly decreased compared to the rest of the test run. This spike in velocity directly translated to the error in the the  $X$ -direction and  $Y$ -direction.

To represent the multidimensional displacement, the Euclidean distance ( $l_2$  Norm) between the predicted  $(\Delta X_p, \Delta Y_p)$  and the truth  $(\Delta X_t, \Delta Y_t)$  is calculated with Equation 3.21. The difference between the displacements are shown in Figure 4.11 and Table 4.3 present the subsequent error distribution statistics for each direction of error distributions respectively.

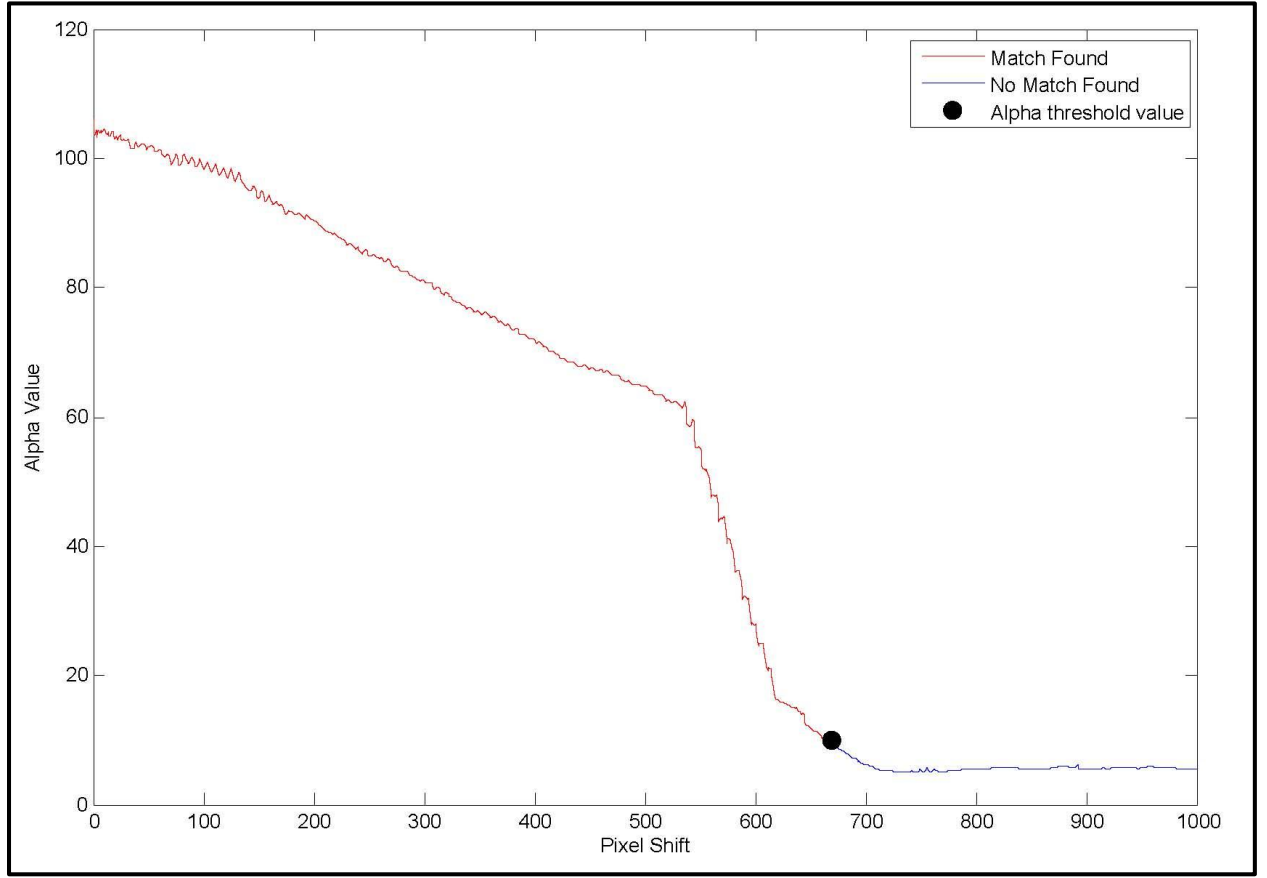


Figure 4.7: Alpha (A) Parameter Sweep Data.

Observe that, consistent with the individual direction assessments, the ( $l_2$  Norm) error hovers around the mean and has a large spike consistent with the spike in velocity. Although this anomaly in the data is consistent with the spike in the velocity there are a number of other sources of possible error. For instance, changes in camera distance from the ground from uneven surfaces will introduce camera calibration error. Also notice the “gaps” in the data, where all data points are zero, these gaps are

Table 4.3: Statistics for Multidimensional Displacement Differences.

Parameter	Value
Mean	1.347 cm
Std Dev	15.268 cm

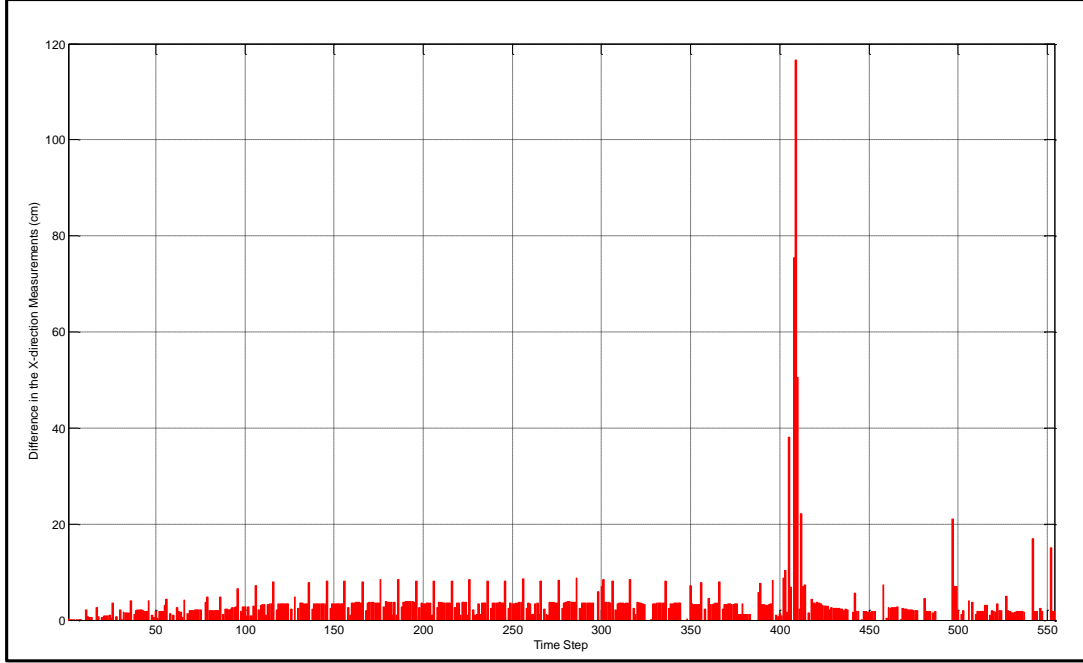


Figure 4.8: Difference Between Predicted X-Displacements and Truth X-Displacements.

due to no image correlation calculation which may be associated with previously mentioned sources of error.

### 4.3 Summary

Chapter IV presented the hardware and testing parameters used to test the validity of the optical flow based registration algorithm. Testing resulted in 2 cm level accuracy with an  $\alpha$  parameter setting of 10. Finally, recommendations for future work and final remarks will conclude this thesis in Chapter V.

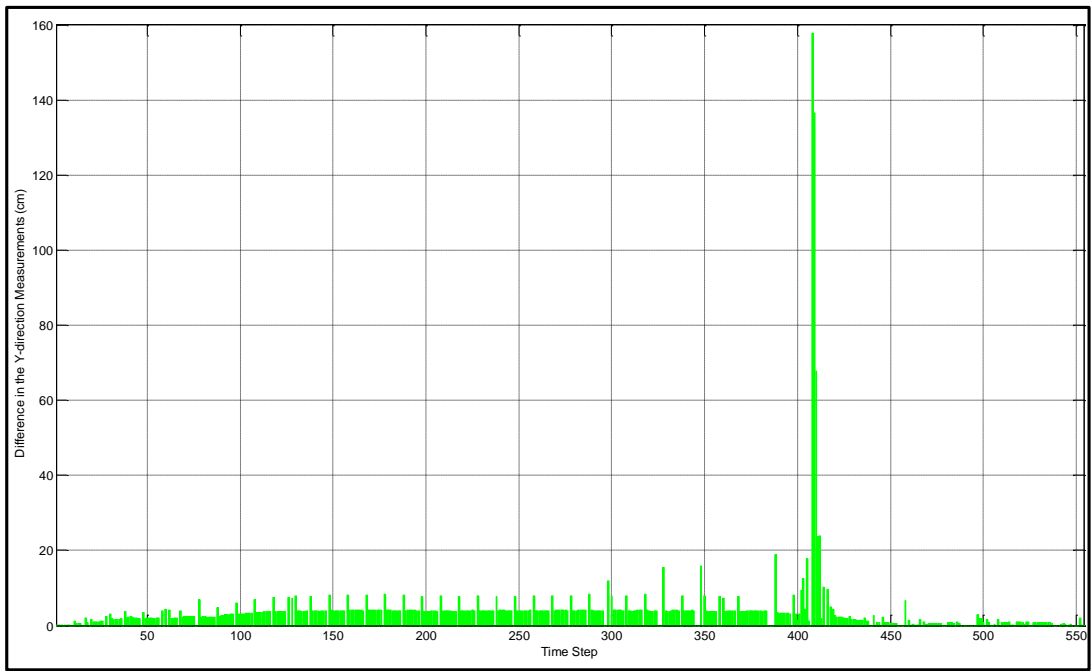


Figure 4.9: Difference Between Predicted Y-Displacements and Truth Y-Displacements.

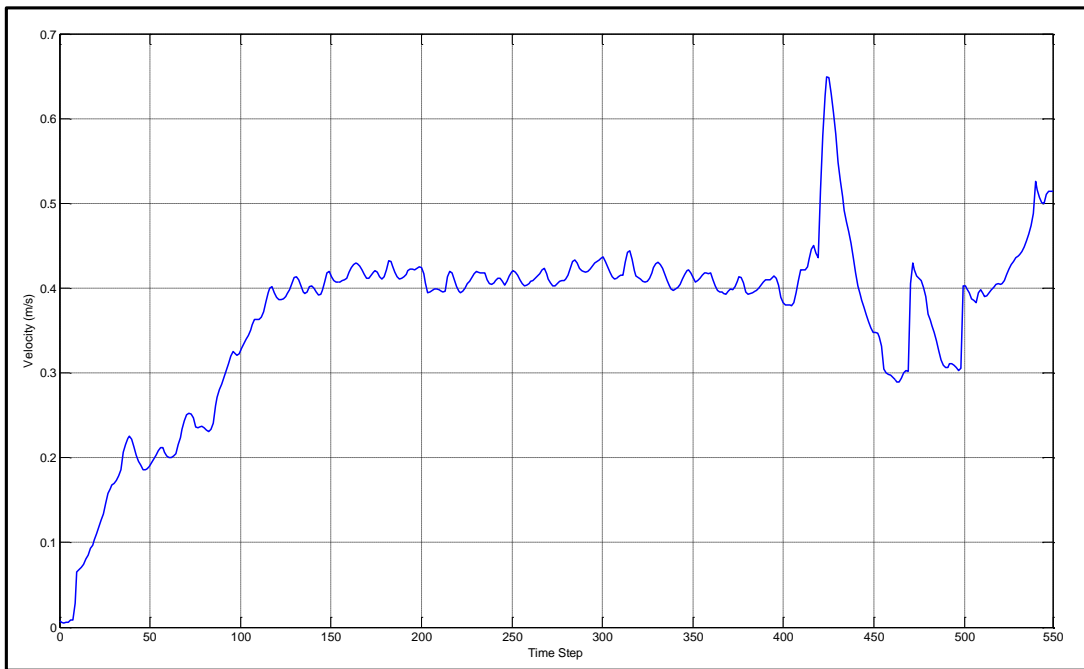


Figure 4.10: Average Velocity Over Test Run

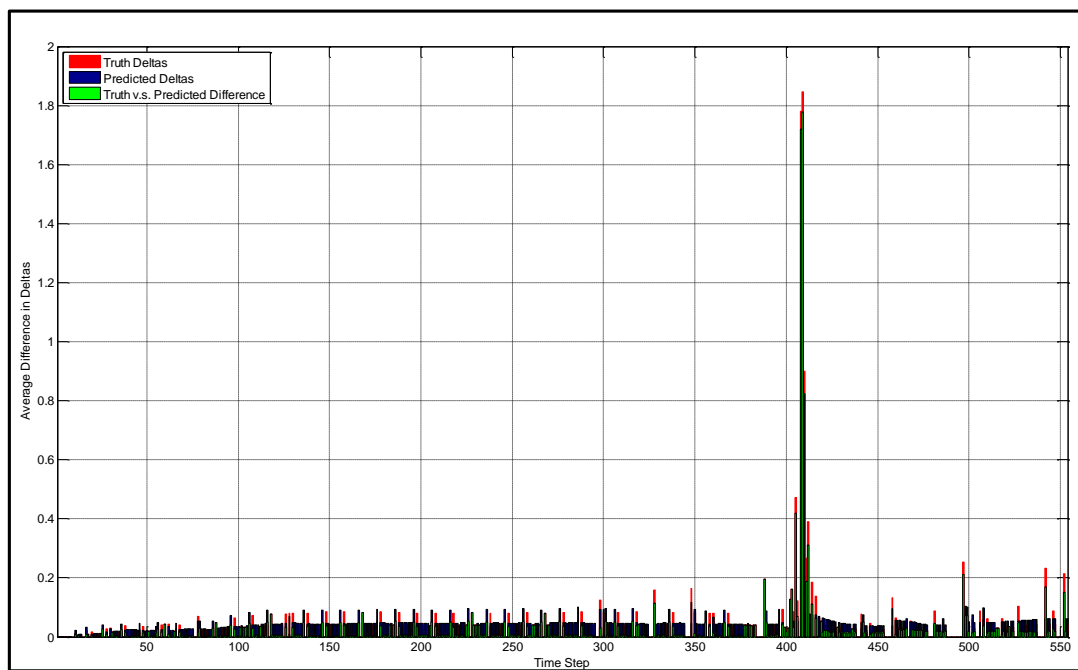


Figure 4.11: Multidimensional Displacement Error



## V. Conclusions

This chapter presents the conclusions of this thesis and recommendations for future work. Section 5.1 presents an evaluation of the algorithm developed. Improvements for the navigation system and recommendations for the next phase in research are presented in Section 5.2.

### 5.1 *Conclusions*

Analysis of the optical flow based phase correlation algorithm began with characterization of a single image presented in Section 3.1.1. This technique proved to be beneficial for initial augmentation of the algorithm because truth data and testing options were readily available and reliable. Next, the algorithm was tested using real-world images from a single camera presented in Section 3.1.2.

The concept of using image registration methods to conduct optical flow odometry is a relatively new technique. By reviewing image registration algorithms analyzed by [11] as well as the results produced in this research, the phase correlation method is “regarded as a robust estimator in the presence of noise” [12].

The system description in Chapter IV outlines a system that captures images at a rate of 10 Hz while the vehicle traverses through an environment. The optical flow odometry algorithm then analyzes the translational motion between images to produce the displacement in vehicle pose over time. These results can be further extended to produce pose estimation.

This research also found that phase correlation methods will result in a displacement estimation regardless of whether or not a specific shift was found between 2 images. Therefore, an alpha parameter was established to detect such false positives. Setting  $\alpha = 10$  yielded no false positives throughout the dataset. This greatly improves the accuracy of the overall technique.

The application of the optical flow based image registration revealed that the spatial pose displacement averaged only 1.3 cm in error throughout the data set.

This shows that further extension to pose estimation would yield similar metric level accuracy. This proves that this process is a viable solution for navigation purposes. While the scope of this research was limited to optical flow odometry, it can be further expanded and characterized for mapping capabilities.

## ***5.2 Recommendations for Future Research***

This section discusses areas for future work which would provide further advancement of the navigation capability presented in this thesis. Investigation into the algorithm response to differing terrain responses, varying image capture rates, and limited lighting conditions that replicate what will be encountered in the underground tunnel setting will further develop the optical flow based registration algorithm for implementation into an autonomous mobile robot.

Errors associated with changes in velocity of vehicle may be able to be reduced or eliminated by using a variable image capture rate that directly correlates with the velocity of the vehicle. The ability to dynamically adjust the image capture rate based on the velocity may significantly decrease the amount of data storage needed and increase accuracy. For example, if the vehicle is traversing down a long straight tunnel at a constant velocity of a capture rate of 5 Hz may be adequate for accuracy cross-correlation of the images but if the vehicle encounters a downhill and the velocity of the vehicle increased a capture rate of 10 Hz may be needed to accurately calculate displacement [34].

Another area for future research dealing with image exposure may pose a significant increase in algorithm functionality. Normalizing the image exposure for each image capture by synchronizing a strobe light that matches the frame rate may minimize the amount of discrepancies between images that do not correlate with actual displacement [34].

Another area for future research is investigation into a solution of uneven surfaces encountered in the underground tunnel environment. Mounting laser range

finders in parallel with each of the cameras could provide the ability to more accurately determine roll, pitch and yaw of the cameras which could potentially allow for camera calibration adjustments on the fly to account for the differences. Camera calibration on the fly could potentially produce more accurate image capture for increased displacement accuracy [34].

To prove the true viability of the optical flow based registration algorithm presented in this thesis is to integration into a Kalman filter to resolve actual path traversed and development of a map. The integration into a Kalman filter would make it possible reconstruct the actual path the vehicle traversed for real-time image processing to update IMU for an accurate position solution.

The ultimate goal of this research is for this navigation technique to be implemented into a full scale rover that is able to accurately and autonomously navigate through an underground tunnel-like environment. This capability meets the immediate needs of users such as the US Homeland Security and our military.

## Bibliography

1. Adams, H., S. Singh, and D. Strelow. "An Empirical Comparison of Methods for Image-based Motion Estimation". In *Proceedings of the 2002 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, Volume 1, 123128. Sept 2002.
2. Amidi, O., T. Kanade, and J. Miller. "Vision-Based Autonomous Helicopter Research at CMU". *Heli Japan* 98. 1998.
3. Arras, K. *Feature-Based Robot Navigation in Known and Unknown Environments*. Ph.D. Dissertation, Swiss Federal Institute of Technology, 2003.
4. Barnea, D. and H. Silverman. "A Class of Algorithms for Fast Digital Registration". *IEEE Trans. Comput.* C-21:179–186, 1972.
5. Beauchemin, S. and J. Parron. "The Computational of Optical Flow". *ACM Computing Surveys*, 27(3):433–367, 1995.
6. Beliveau, Y., J. Fithian, and M. Deisenroth. "Autonomous Vehicle Navigation With Real-Time 3D Laser Based Positioning for Construction". *Automation in Construction*, 5(4):261–272, 1996. ISSN 0926-5805.
7. Bonarini, A., M. Matteucci, and M. Restelli. "Automatic Error Detection and Reduction for an Odometric Sensor Based on Two Optical Mice". *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 1675–1680. Barcelona, Spain, Apr 2005.
8. Borenstein, J. "Experimental Results From Inertial Odometry Error Correction With the OmniMate Mobile Robot". *IEEE Transactions on Robotics and Automation*, 14:963–969, 1998.
9. Borenstein, J. and L. Feng. "Measurement and Correction of Systematic Odometry Errors in Mobile Robots". *IEEE Transactions on Robotics and Automation*, 12:869–880, 1996.
10. Britting, K. *Inertial Navigation System Analysis*. Wiley Interscience, New York, 1971.
11. Brown, L. "A Survey of Image Registration Techniques". *ACM Computing Surveys*, 24(4):325–376, Dec 1992.
12. Cain, S., M. Hayat, and E. Armstrong. "Projection-Based Image Registration in the Presence of Fixed-Pattern Noise". *IEEE Transactions on Image Processing*, 10:1860–1872, Dec 2001.
13. Campbell, J., R. Sukthankar, and I. Nourbakhah. "Techniques for Evaluating Optical Flow for Vision Odometry in Extreme Terrain". *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference*, Volume 4, 3704–3711. 28 2004.

14. Cracknell, A. and L. Hayes. *Introduction to Remote Sensing*. CRC Press, 2nd Edition, 2007.
15. Gennery, D. and H. Moravec. *Cart Projection Progress Report*. Master's Thesis, Stanford University, contract NASW, 1976.
16. Giachetti, A., M. Campani, and V. Torre. "The Use of Optical Flow for Road Navigation". *IEEE Transactions on Robotics and Automation*, 14(1):34–48, Feb 1998.
17. Gonzalez, R., R. Woods, and S. Eddins. *Digital Image Processing Using MATLAB*. Gatesmark Publishing, 2009.
18. Holstege, S. "More drug tunnels being found on border". *The Arizona Republic*, Nov 2008.
19. Hong, P. and K. Hyun. *Optical Flow Through Mechanical Eyes*. Technical Report 209, Information and Communications University, South Korea, Aug 2008.
20. Hrabar, S. and G. Sukhatme. "A Comparison of Two Camera Configurations for Optic-Flow Based Navigation of a UAV Through Urban Canyons". In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Volume 3, 26732680. Sept 2004.
21. Kim, T. and J. Lyou. "Indoor Navigation of Skid Steering Mobile Robot Using Ceiling Landmarks". In *Proceedings of the 2009 IEEE International Symposium on Industrial Electronics*. Seoul, Korea, July 2009.
22. Konolige, K., J. Augenbraun, N. Donaldson, C. Fiebig, and P. Shah. "A Low-cost Laser Distance Sensor". *IEEE International Conference on Robotics and Automation*, 3002–3008. Pasadena, CA, May 2008.
23. Kuglin, C. and D. Hines. "The Phase Correlation Image Alignment Method". 163–165. New York, NY, 1975.
24. Kumar, V. and M. Nanda. *Image Processing in Frequency Domain Using Matlab: A Study For Beginners*. Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan-173 215, India, 1st Edition, Sept 2008.
25. Laviers, K. *Concurrent Cognitive Mapping and Localization Using Expected Maximization*. Master's Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 2004.
26. Lee, M., R. Marks, S. Rock, and R. Marks. "Automatic Visual Station Keeping of an Underwater Robot". *Proceedings of IEE Oceans*, 2:137–142, 1994.
27. Leonard, J. and H. Durrant-Whyte. "Mobile Robot Localization by Tracking Geometric Beacons". *Robotics and Automation, IEEE Transactions on*, 7(3):376–382, Jun 1991. ISSN 1042-296X.

28. Lomanno, C. and K. Fisher. “Computer Mouse-Based Odometry and Heading for Indoor Navigation”. *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Insititure of Navigation*, 2262–2270. Portland, OR, Sept 2010.
29. McElroy, J. *Navigation Using Signals of Opportunity in the AM Transmission Band*. Master’s Thesis, Air Force Institute of Technology, 2006.
30. Nowak, R. “Intro to Digital Signal Processing”. *Common Creative Attribution* , 7–10, 2004.
31. Optics, VS Techology Machine Vision. “Lenses for Machine Vision”, Jun 2010. URL <http://ww.vst.co.jp/index.php>.
32. PixeLINK. “PixeLINK PL-A741 Machine Vision Camera Datasheet”. Apr 2004. URL <http://www.pixelink.com/>.
33. Press, W., W. Vetterling, S. Teukolsky, and B. Flannery. *Numerical Recipes in C++: The art of scientific computing*. Cambridge University Press, New York, NY, 2nd Edition, 2002.
34. Raddin, J. “Solutions to the Underground Navigation Problem”, Mar 2011.
35. Reed, R. *Comparison of Subpixel Phase Correlation Methods for Image Registration*. Technical Report ADA519383, USAF Air Force Material Command (Arnold Engineering Development Center), Arnold AFB, TN, Apr 2010.
36. Roberts, R., H. Nguyen, N. Krishnamurthi, and T. Balch. “Memory-Based Learning for Visual Odometry”. *IEEE International Conference on Robotics and Automation*, 47–52. Pasadena, CA, May 2008.
37. Schempf, H., W. Crowley, C. Gasior, and D. Moreau. “Ultra-rugged Soldier-Robot for Urban Conflict Missions”. *Systems International (AUVSI) 30th Annu. Symp. and Exhibition*. 2003.
38. Titterton, D. and J. Weston. *Strapdown Inertial Navigation Technology*. The American Institute of Aeronautics and Asteronautics, 2nd Edition, 2004.
39. Tzimiropoulos, G. and T. Stathaki. “Estimation of Subpixel Shifts Using Gradient Correlation”. *5th SEAS DTC Technical Conference*. Edinburg, 2010.
40. Veth, M. *Fusion of Imagery and Intertial Sensors for Navigation*. Ph.D. Dissertation, Air Force Institute of Technology, Wright-Patterson AFB, OH, Sept 2006.
41. Zheng, G., E Smith, N. Rafla, and T. Andersen. “2X1D Image Registration and Comparison”. *Proc. SPIE Electronic Imaging-Image Processing: Machine Vision Applications III*, Volume 7538. San Jose, CA, Jan 2010.

# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE</b> (DD-MM-YYYY) 24-03-2011		<b>2. REPORT TYPE</b> Master's Thesis		<b>3. DATES COVERED</b> (From — To) Sep 2010 — Mar 2011	
<b>4. TITLE AND SUBTITLE</b>  Optical Flow-Based Odometry for Underground Tunnel Exploration				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b>	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b>  Terra Kier, 1st Lt USAF				<b>5d. PROJECT NUMBER</b> N/A	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  AFIT/GE/ENG/11-21	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Air Base Technologies Division (Brian Skibba) 139 Barnes Drive, Suite 2 Tyndall Air Force Base, FL 32403 (850-283-3725 brian.skibba@tyndall.af.mil)				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> AFRL/RXQF	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>  As military operations in degraded or GPS-denied environments continue to increase in frequency and importance, there is an increased necessity to be able to determine precision location within these environments. Furthermore, authorities are finding a record number of tunnels along the U.S.-Mexico border; therefore, underground tunnel characterization is becoming a high priority for U.S. Homeland Security as well. This thesis investigates the performance of a new image registration technique based on a two camera optical-flow configuration using phase correlation techniques. These techniques differ from other image based navigation methods but present a viable alternative increasing autonomy and answering the tunnel based navigation problem. This research presents an optical flow based image registration technique and validates its use with experimental results on a mobile vehicle. Results reveal that further extension to pose estimation yields accuracy within 1.3 cm.					
<b>15. SUBJECT TERMS</b>  Optical-Flow Odometry; Image Registration; Cross-correlation via Fourier Analysis; Underground Navigation; Image Processing; Mobile Navigation; Vehicle Localization; Robotic Mapping					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>  UU	<b>18. NUMBER OF PAGES</b>  63	<b>19a. NAME OF RESPONSIBLE PERSON</b> Maj Kenneth A. Fisher, ENG
<b>a. REPORT</b>  U	<b>b. ABSTRACT</b>  U	<b>c. THIS PAGE</b>  U			<b>19b. TELEPHONE NUMBER</b> (include area code) (937) 255-3636, ext 4677; kenneth.fisher@afit.edu